KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 Eye Street, NW
Suite 1100
Washington, DC 20006

October 22, 2012

John J. O'Brien
Director
Healthcare and Insurance
Office of Personnel Management
c/o MSPP@opm.gov

Director O'Brien:

The Center for Democracy & Technology ("CDT") is a non-profit Internet and technology advocacy organization that promotes public policies that preserve privacy and enhance civil liberties in the digital age. As information technology is increasingly used to support the exchange of medical records and other health information, CDT, through its Health Privacy Project, champions comprehensive privacy and security policies to protect health data. CDT promotes its positions through public policy advocacy, public education, and litigation, as well as through the development of industry best practices and technology standards. Recognizing that a networked health care system can lead to improved health care quality, reduced costs, and empowered consumers, CDT is using its experience to shape workable privacy solutions for a health care system characterized by electronic health information exchange.

CDT is frequently relied on for sound policy advice regarding the challenges to health privacy and security presented by health information technology (health IT) initiatives. We have testified before the U.S. Congress five times since 2008 on the privacy and security issues raised by health IT, and we chair the privacy and security policy working group of the federal Health IT Policy Committee.

CDT respectfully submits these comments in response to the U.S. Office of Personnel Management's Multi-State Plan Program (MSP) Application Draft, issued on September 21, 2011.

We commend OPM for suggesting several important privacy features in this draft. We are pleased to see that OPM will evaluate MSP candidates on their privacy and security compliance, as listed in 5) under "Utilization/quality assurance." We also commend OPM for requiring applicants to describe their compliance with Fair Information Practice Principles, listed in 5) under "IT Systems, security and confidentiality." We urge OPM to retain these evaluation criteria in the final MSP application.

However, we have some recommendations to improve the section "IT Systems, security and confidentiality." In particular, we are concerned about the requirement for applicants to transmit line-level data claims to OPM and to describe their system infrastructure's capacity to securely interface with OPM for data transfers, including enrollment, reconciliation, claims/encounter data, and reports (p. 15).

We appreciate OPM's interest in routinely analyzing line-level plan data; effectively managing the MSP program depends on access to data that will be needed for the defined set of purposes described in rules and guidance for all Exchange health plans. However, we believe that OPM's plan to centrally collect copies of this data creates unnecessary privacy and security risks.

A centralized system increases the risk and severity of data breaches by unnecessarily duplicating data and centrally collecting it in one location. Even when the data that is collected is de-identified, there remains some risk of breach and re-identification. Requiring personal data to be shared with government agencies (when such sharing is unnecessary) has the potential to erode public trust in health reform initiatives.[1]

Consequently, we urge OPM to meet its data analytic needs using a decentralized approach, which enables robust data analysis but typically leaves data housed with its original source. There are two general approaches to a decentralized system: the "distributed access" approach and the "distributed query" approach.[2] The best fit for OPM will depend on OPM's particular analytic needs and resource constraints.

Under the distributed access model, MSPs would provide OPM with access to structured data in a secure environment, such as on an edge server or in a cloud storage environment. MSPs will need to populate the edge server or cloud with data formatted in accordance with OPM standards. OPM can securely access the data to meet their analytic needs, retaining the results of their analyses but leaving the original data on the server or in the cloud. This model enables data to be accessed at reduced cost to OPM and with less risk of data breach. As noted in more detail below, the Centers for Medicare and Medicaid Services (CMS) is using this approach to perform the risk adjustment required by the Affordable Care Act.

---

[1] See Rep. Tim Huelskamp, *Obamacare HHS rule would give government everybody's health records*, 23 September 2011, http://washingtonexaminer.com/opinion/op-eds/2011/09/obamare-hhs-rule-wouldgive-government-everybody-s-health-records. See also Rep. Denny Rehberg, *Chairman Rehberg Investigates Possible Violations of Private Health Care Information Under President Obama's Health Care Plan*, 13 October 2011, http://pressrehberg.congressnewsletter.net/mail/util.cfm?gpiv=2100078808.1461.269&gen=1.

[2] For more details about decentralized approaches to health data analytics, see Center for Democracy and Technology, *Decentralizing the Analysis of Health Data*, 22 March 2012, https://www.cdt.org/files/pdfs/Decentralizing-Analysis-Health-Data.pdf.

Another approach is the distributed query model. In one implementation model for this approach, OPM would write the analytic code and send the code to the MSPs; the MSPs then analyze their in-house data using the code and provide the results to OPM. This model also requires use of a common data format; risk of fraud or inaccurate results is reduced if the MSPs are prohibited from writing or manipulating the analytic code. The Food and Drug Administration's Mini-Sentinel Initiative is an example of this approach.

Another implementation model would allow the MSPs to write their own analytic code to address questions from OPM; such a model may be useful in circumstances where there is little risk of fraud or the likelihood (or consequences) of inaccurate results are low. The Federal Partners project of the Food and Drug Administration is an example of this system.

We initially voiced concerns about centralized data collection with respect to OPM's Health Claims Data Warehouse (HCDW), created in 2010 and designed to contain copies of detailed electronic health records of millions of Americans insured by the Federal Employees Health Benefits Program.. In response to letters from CDT and other privacy groups, OPM made improvements to this database in 2011. For example, OPM committed to full compliance with HIPAA and FISMA protections, explained specifically what data would be collected and how it would be used, and pledged to use only de-identified information for analysis purposes and release to external parties.

However, the HCDW system is still centralized. It is not clear why OPM continues to use centralized model when a decentralized model would minimize data transfer and reduce the risk of data breach, minimize costs, and remain in line with public expectations of privacy. It also is not clear why OPM does not firmly commit to using cryptographic techniques to maintain data security, such as one-way hash function, which would scramble identifiers in the health information but still allow for data comparison.

We continue to urge OPM to pursue a decentralized approach to meet its data analysis needs. The decentralized approach should also explicitly require that any claims data accessed by OPM be de-identified according to HIPAA standards. OPM's centralized data collection approach is directly contrary to the approach being taken by other federal agencies seeking to perform real-time "big data" analytics. For example, in July 2011, the Centers for Medicare and Medicaid Services (CMS) proposed a rule that would have required every state (or HHS on the state's behalf) to collect claims data from every payer in the individual and small group market to support the risk adjustment program mandated by the Affordable Care Act. After considering the decentralized approach recommended by CDT and other organizations, CMS opted to change the database to a distributed access model in March 2012. Their approach follows CDT's recommendations for a successful decentralized model.

Another example is the Food and Drug Administration's (FDA) Mini-Sentinel Initiative. Sentinel was launched in 2008 in order to quickly monitor the safety of products the FDA regulates. Mini-Sentinel is a distributed query model that

provides a secure web interface through which users authorized by the FDA can query product data and send questions to the data sources (which include health plans). The data remain with and are managed by the participating data sources, reducing the risk of fraud and inaccuracies by the data sources.

Unfortunately, there is still a general trend among some businesses and government agencies to develop a new database for every analytic need, and OPM is continuing to follow this outdated and privacy-risky model. Although CDT supports cost-cutting and fraud detection goals of health claims databases, individual privacy and data security are ill served when repositories and copies of identifiable personal information are created unnecessarily. To the extent possible, government agencies and businesses should seek to meet their objectives through methods that leverage existing systems, minimize data transfer, and maintain the relative anonymity of data subjects.

Thank you for the opportunity to comment on the application.

Sincerely,

Deven McGraw
Director, Health Privacy Project

Cc: Meredith Whipple, Policy Analyst