Bloomberg BNA

Privacy and Security Law Report®

HIPAA Business Associates and Health-Care Big Data: Big Promise, Little Guidance

BY REECE HIRSCH AND HEATHER DEIXLER

Health-care "big data" projects hold the promise of transforming health-care delivery by permitting providers to assess acute cases within the context of an entire population, developing new methods of identifying and preventing illness, enabling new discoveries and reducing health-care costs. However, unlike less regulated industries such as retail, the utilization of big data in health care must often be conducted within the parameters of rigorous, industry-specific privacy laws and regulations—most notably the Health Insurance Portability and Accountability Act (HIPAA). In order to realize the promise of health-care big data, companies must navigate some of the less well-defined aspects of HIPAA.

The health-care industry remains to a large degree local and regional, rather than national, causing much health-care data to be siloed and utilized primarily by the entity that created it. While hospitals, physicians and most health plans operate within a region or state, there are many vendors to the health-care industry that are nationwide in scope, assembling vast databases of individually identifiable health information from their provider and plan customers. These companies offer a wide variety of services, including electronic health records (EHRs) , cloud-based software and outsourcing, coding and billing, pharmaceutical benefit management, pharmaceutical distribution and claims processing and administration.

Reece Hirsch is a partner in the San Francisco office of Morgan, Lewis & Bockius LLP. He can be reached at (415) 442-1422 or rhirsch@morganlewis.com.

Heather Deixler is an associate in the San Francisco office of Morgan Lewis. She can be reached at (415) 442-1422 or hdeixler@morganlewis.com.

A common denominator among many of these vendors to the health-care industry is that they are business associates within the meaning of HIPAA. As business associates, these vendors are generally prohibited from using protected health information (PHI) for purposes other than providing services to a HIPAA-covered entity in accordance with the terms of a services agreement. HIPAA does, however, provide some limited latitude for business associates to use and disclose PHI for other purposes. This article will examine some of the fuzzy rules governing use and disclosure of PHI by business associates and consider the extent to which they enable or pose obstacles to health-care big data projects.

What Is Big Data?

The term big data typically refers to the application of emerging techniques in data analytics, such as machine learning and other artificial intelligence tools, to enormous stores of personal information. Individually identifiable data are being assembled in ever larger and more comprehensive databases, from diverse sources such as personal health records, medical records, claims data, Web-browsing data trails, GPS devices, social networking activity and biometric sensor data. The term big data refers to the powerful and often surprisingly granular information that can be assembled about individuals based upon analysis of these enormous databases. Yahoo! Inc. Chief Executive Officer Marissa Mayer vividly described big data as "watching the planet develop a nervous system."[1]

A December 2013 report by the Bipartisan Policy Center, based on a policy forum held earlier last year, considered the potential of big data in health care and identified privacy and security as an often-cited barrier to progress on the use and exchange of big data. The report noted, "While . . . HIPAA is designed to safe-

[1] Daniela Hernandez, *Big Data Is Transforming Healthcare*, WIRED.com, Oct. 16, 2012, http://www.wired.com/wiredscience/2012/10/big-data-is-transforming-healthcare/.

guard patient privacy, it is often misunderstood, misapplied and over-applied in ways that may inhibit information sharing unnecessarily.''[2]

The Bipartisan Policy Center is absolutely right that HIPAA is often misunderstood when it intersects with big data analytics. Unfortunately, many of those misunderstandings are inevitable because the rules governing data mining and analysis by business associates are not always clear.

Management and Administration by Business Associates

Pursuant to the mandated terms of a business associate agreement, a business associate is prohibited from using or further disclosing the covered entity's PHI other than as permitted or required by the business associate agreement or as required by law.[3] A business associate agreement between a business associate and a covered entity may permit the business associate to ***use*** the information received by the business associate in its capacity as a business associate, if necessary ''[f]or the proper management and administration of the business associate . . . .''[4]

A business associate agreement may also permit a business associate to ***disclose*** information received by the business associate in its capacity as a business associate for its management and administration purposes if:

> (A) The disclosure is required by law; or
>
> (B)(1) The business associate obtains reasonable assurances from the person to whom the information is disclosed that it will be held confidentially or used or further disclosed only as required by law or for the purpose for which it was disclosed to the person; and
>
> (2) The person notifies the business associate of any instances of which it is aware in which the confidentiality of the information has been breached.[5]

Most business associate agreements contain provisions permitting the business associate to use and disclose PHI for its ''proper management and administration'' purposes, in accordance with the provisions cited above. In the absence of such a provision, a business associate might not be permitted to use PHI for many activities vital to conducting its business. Business associates managing large volumes of PHI often must parse this ''management and administration'' rule to determine the extent to which it permits certain uses of big data.

The terms ''management'' and ''administration'' are not expressly defined under the HIPAA Privacy Rule, and there is a surprising lack of guidance or commentary from the Department of Health and Human Services (HHS) Office for Civil Rights (OCR) elaborating on the intended meaning of those terms. The most relevant supplemental guidance is found in the following commentary to the proposed Privacy Rule:

> *Comment:* A commenter recommended that the business partner contract specifically address the issue of data mining because of its increasing prevalence within and outside the health care industry.
>
> *Response:* We agree that protected health information should only be used by business associates for the purposes identified in the business associate contract. We address the issue of data mining by requiring that the business associate contract explicitly identify the uses or disclosures that the business associate is permitted to make with the protected health information. Aside from disclosures for data aggregation and ***business associate management***, the business associate contract cannot authorize any uses or disclosures that the covered entity itself cannot make. Therefore, data mining by the business associate for any purpose not specified in the contract is a violation of the contract and grounds for termination of the contract by the covered entity.[6]

This commentary from the HHS makes clear that, while uses and disclosures of PHI for ''business associate management'' purposes may be permissible, a business associate may not engage in data mining of PHI if it is not permitted by the business associate agreement and it is inconsistent with the covered entity's HIPAA obligations. The commentary also could be read to suggest that a business associate is prohibited from using PHI for the business associate's commercial purposes unrelated to the services that a covered entity has contracted for and not expressly authorized by a business associate agreement, such as data mining.

It seems reasonable to characterize certain activities as relating to the management and administration of the business associate's business, and thus permitted under the terms of its business associate agreements, including use of PHI for: (1) quality assurance, (2) utilization review, (3) compliance, (4) fraud prevention, (5) auditing and (6) cost-management and planning-related analyses. HIPAA-covered entities are permitted to engage in similar uses of PHI as part of their ''health care operations'' activities. It could be reasonably asserted that these activities are integral to a business associate's current and future suite of products and services. Such activities could also be characterized as internal ''back office'' activities that permit the business associate to more effectively utilize the PHI of a client for the benefit of that client.

But how should a business associate interpret these rules when effective management of its business ***requires*** data mining? What if data mining of customer data is necessary in order to develop the next iteration of the business associate's product or service? What if crawling and mapping of customer data is necessary in order to facilitate the provision of future or anticipated products or services? What if the business associate must engage in data mining as part of its research in order to identify and develop its next offering? These uses of big data are not strictly necessary in order for the business associate to provide the contracted service to a HIPAA-covered entity, but they may very well be critical to management and administration of the business associate's enterprise and providing value to customers through improved products and services.

In the absence of interpretive guidance from the OCR on the meaning of ''management and administration,'' a business associate must rely almost entirely on the

[2] Bipartisan Policy Ctr., *A Policy Forum on the Use of Big Data in Health Care* 6 (Dec. 2013), *available at* http://bipartisanpolicy.org/sites/default/files/Use%20of%20Big%20Data%20in%20Health%20Care.pdf.

[3] 45 C.F.R. § 164.504(e)(2)(ii)(A).

[4] *Id.* § 164.504(e)(4)(i)(A).

[5] *Id.* § 164.504(e)(4)(ii).

[6] Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,462, 82,644 (Dec. 28, 2000) (emphasis added).

plain meaning of those terms, which are open to interpretation. A business associate might mitigate the risks associated with uncertain reliance on the ''management and administration'' provision by obtaining express or implied consent from its clients with respect to data analytics functions. If the business associate's agreements with clients expressly authorize the business associate to perform big data analysis as part of its management and administration activities, then the risks associated with the practice would be reduced.

Data Aggregation by Business Associates

HIPAA's Privacy Rule permits business associates to perform data aggregation services relating to the ''health care operations'' of the covered entity from which it receives the information. ''Data aggregation'' is defined as a business associate's combining of PHI received from multiple covered entities ''to permit data analyses that relate to the *health care operations* of the respective covered entities.''[7]

In its commentary to the final HIPAA Privacy Rule, the HHS explained that it included data aggregation services as a permitted provision in business associate agreements to ''clarify the ability of covered entities to contract with business associates to undertake quality assurance and comparative analyses that involve the protected health information of more than one contracting covered entity.''[8] The HHS further noted:

> We except data aggregation from the general requirement that a business associate contract may not authorize a business associate to use or further disclose protected health information in a manner that would violate the requirements of this subpart if done by the covered entity in order to permit the combining or aggregation of protected health information received in its capacity as a business associate of different covered entities when it is performing this service. In many cases, the combining of this information for the respective health care operations of the covered entities is not something that the covered entities could do—a covered entity cannot generally disclose protected health information to another covered entity for the disclosing covered entity's health care operations. ***However, we permit covered entities that enter into business associate contracts with a business associate for data aggregation to permit the business associate to combine or aggregate the protected health information they disclose to the business associate for their respective health care operations***.[9]

The phrase ''health care operations'' is broadly defined to include a laundry list of activities, including the following:

> (1) Conducting quality assessment and improvement activities, including outcomes evaluation and development of clinical guidelines, provided that the obtaining of generalizable knowledge is not the primary purpose of any studies resulting from such activities; ***population-based activities relating to improving health or reducing health care costs***, protocol development, case management and care coordination, contacting of health care providers and patients with information about treatment alternatives; and related functions that do not include treatment;[10]

Because ''health care operations'' include ''population-based activities relating to improving health or reducing health care costs,'' business associate data aggregation services may be readily aligned with the population-based health objectives of many big data projects.

Because ''health care operations'' include ''population-based activities relating to improving health or reducing health care costs,'' business associate data aggregation services may be readily aligned with the population-based health objectives of many big data projects.

In summary, business associates may utilize big data in providing data aggregation services to their customers, provided that: (i) the business associate enters into business associate agreements that permit data aggregation services; (ii) all of the PHI analyzed/utilized by the business associate as part of its data aggregation service is received by the business associate in its capacity as a business associate of a HIPAA-covered entity; (iii) the business associate customers receiving the product of the data aggregation services are covered entities for which the business associate is acting as a business associate; and (iv) the data aggregation services relate to one of the types of activities enumerated in the definition of ''health care operations.'' Although HIPAA's definition of ''data aggregation'' is relatively broad (*i.e.*, the combination of PHI by a business associate from multiple covered entities ''to permit data analyses that relate to the health care operations of the respective covered entities''), the OCR has not provided any detailed guidance or commentary regarding the scope of the activities that may constitute ''data aggregation'' services.

For instance, data aggregation could be useful for a pharmacy benefit manager (PBM) to identify drug utilization trends by aggregating the data obtained from its covered entity clients. Since ''health care operations'' includes ''population-based activities relating to improving health or reducing health care costs'' and ''care coordination,'' a PBM could combine the PHI it receives from multiple covered entities in order to extract data, perform the analytics and provide the appropriate covered entity with the results of the analysis, provided that the PBM's business associate agreements with its covered entity clients permit data aggregation activities.

If a business associate combines the data of multiple types of HIPAA-covered entities, such as health-care providers and health plans, in performing data aggregation services, then the business associate should consider whether the resulting data analysis relates to the health-care operations activities of each covered entity receiving that analysis. For example, the health-care operations of health-care providers and health plans differ significantly. Some categories within the definition of ''health care operations'' are applicable only to health plans, such as ''underwriting, enrollment and premium rating.'' Categories such as population-based

[7] 45 C.F.R. § 164.501 (definition of ''Data aggregation'') (emphasis added).

[8] 65 Fed. Reg. at 82505.

[9] *Id.* (emphasis added).

[10] 45 C.F.R. § 164.501.

health activities are generally applicable to both health-care providers and health plans.

It is important to remember that data analysis performed under the data aggregation exception may only be shared with the covered entities that shared the PHI with the business associate. However, if a business associate also has permission to de-identify PHI under the terms of a business associate agreement, then the analysis performed through data aggregation may meet HIPAA's de-identification standard, in which case it may be shared with any third party. That brings us to the last of the three important rules governing the use of big data by business associates.

De-identification by Business Associates

Health information that does not identify an individual, and where there is no reasonable basis to believe that the information can be used to identify an individual, ceases to be PHI and is deemed to be ''de-identified.''[11] The recent advancement of health information technologies enabling companies to capture large quantities of health-care data has created the potential to combine these data to conduct comparative effectiveness studies, scientific research and policy assessment. De-identification of PHI is one method enabling business associates to harness the data they have collected. Unlike the management and administration and data aggregation exceptions, the OCR has issued clear and specific guidance on how covered entities may apply the de-identification standard.

There are two methods of de-identifying PHI. The first involves removing the following identifiers of the individual or of relatives, employers or household members of the individual: (1) names; (2) all geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code and their equivalent geocodes, with certain limited exceptions; (3) all elements of dates (except year) for dates directly related to an individual, including birth date, admission date and discharge date; (4) telephone numbers; (5) fax numbers; (6) e-mail addresses; (7) Social Security numbers; (8) medical record numbers; (9) health plan beneficiary numbers; (10) account numbers; (11) certificate/license numbers; (12) vehicle identifiers and serial numbers, including license plate numbers; (13) device identifiers and serial numbers; (14) ''Web Universal Resource Locators'' (URLs) ; (15) Internet protocol address numbers; (16) biometric identifiers, including finger and voice prints; (17) full face photographic images and any comparable images; and (18) any other unique identifying number, characteristic or code.[12]

In addition to removing these 18 identifiers, the covered entity must also not have actual knowledge that the information could be used alone or in combination with other information to identify an individual who is a subject of the information.[13] The Privacy Rule's standard for de-identification of PHI is quite rigorous, and such de-identified data is generally not useful for analytics intended to target or tailor a product or service to an individual.

The second method of de-identification permits a covered entity to determine that data are not individually identifiable information if a person with appropriate experience with generally accepted statistical and scientific principles and methods of de-identification (1) determines that ''the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and (2) [d]ocuments the methods and results of the analysis that justify such determination.''[14] A potential benefit of this second method of de-identification is that a business associate might be able to retain certain of the 18 identifiers in a data set and still consider the information to be de-identified, assuming the statistician can make the required determination. In November 2012, the OCR published additional guidance on de-identification methods.[15] In defining who was an ''expert'' for purposes of rendering health information de-identified, the OCR noted that no specific professional degree or certification is required, and relevant expertise may be gained through various routes of education and experience.[16] the OCR did note that such experts would typically be found in the statistical, mathematical or scientific domains.[17]

The Privacy Rule's standard for de-identification of protected health information is quite rigorous, and such de-identified data is generally not useful for analytics intended to target or tailor a product or service to an individual.

A covered entity may assign a code or other means of record identification to allow ''re-identification'' by the covered entity once PHI has been de-identified using one of the two methods described above. However, the code or other record identification must not be derived from information about the individual (*e.g.*, using selected digits from a Social Security number), and the covered entity must not use or disclose the code or the mechanism for re-identification.[18] For example, a business associate might, at the direction of a hospital customer, generate a unique code for each patient that the hospital could use to re-identify the data after they have been de-identified. While the de-identified information is no longer considered PHI, the code is considered PHI because it can be used to identify the patient. Thus any disclosure of the code must itself fit within an exception under HIPAA.

As a business associate of its covered entity customers, a business associate may de-identify PHI only if expressly permitted to do so by the terms of its business

[11] *Id.* § 164.514(a).

[12] *Id.* § 164.514(b)(2).

[13] *Id.* § 164.514(b)(2)(ii).

[14] *Id.* § 164.514(b)(1)(i).

[15] OCR, ''*Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*'' (Nov. 26, 2012), *available at* http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf (11 PVLR 1720, 12/3/12).

[16] *Id.*

[17] *Id.*

[18] 45 C.F.R. § 164.514(c).

associate agreement. Although de-identified PHI is generally not very useful in providing patient-specific products and services, one advantage of de-identification from the business associate's perspective is that, unlike data used in data aggregation services, de-identified PHI may be used by the business associate for any purpose because it is no longer considered PHI. Thus, in the above example, a PBM would have greater flexibility in the breadth of analytics it conducted if its customer data were de-identified, and the PBM would not need to ensure that the data were used for purposes of the covered entities' health-care operations.

Takeaways

Large business associate entities are destined to play a key role in the development of big data analytics in health care, but they must operate within the parameters of the often ambiguous HIPAA rules governing uses of PHI for management and administration, data aggregation services and de-identification. Because companies are increasingly basing new products and business models on the use of big data, it is vital that those companies address these issues in customer services agreements prior to the collection of data if possible.

HIPAA-covered entities should be aware of the uses of PHI that are possible under the often-overlooked business associate agreement provisions relating to management and administration, data aggregation and de-identification. Perhaps most significantly given the promise of big data analytics to improve the quality and efficiency of health-care delivery, the OCR should consider providing guidance to clarify the regulatory landscape surrounding this important topic.