wiley

NEWSLETTER

Big Data and Health Care

August 2017

This article is excerpted from a presentation by Kirk J. Nahra (along with Kim Gray of QuintilesIMS) at the American Health Lawyers Association Annual Meeting in San Francisco.

While the big data revolution is having an impact across more and more places throughout our economic system, nowhere are the opportunities and challenges more stark than in the health care industry. Beyond the "normal" ethical and legal challenges that apply in general to big data analytics, the health care industry presents a particularly tricky balance of regulatory ambiguity, business, and strategic overlaps and operational challenges for any business operating in the health care industry or a broader range of companies creating and using health care information (or data that may be useful for health care purposes).

In addition, because of the overall opportunities presented by data for the health care industry, and the particular sensitivities across our regulatory and behavioral structures about health care information, the challenges presented by privacy and security in a data driven health care world may create one of the most important challenges we face in the health care system. How we resolve these issues will impact the success of our health care system in significant ways going forward.

Big Data

So, what is "big data?" Big data can mean simply everything, all data being generated in an increasingly electronic and digital environment. Big data is arriving from multiple new sources (e.g., the Internet of Things) at amazing velocities, volumes and varieties. According to the Federal Trade Commission (FTC), "big data" refers

Practice Areas

Privacy, Cyber & Data Governance

to "a confluence of factors, including the nearly ubiquitous collection of consumer data from a variety of sources, the plummeting cost of data storage, and powerful new capabilities to analyze data to draw connections and make inferences and predictions."

With this idea, to extract meaningful value from big data, you need optimal processing power, storage, analytics capabilities, and data skills. In addition, from a privacy and regulatory perspective, it is important to understand that the people collecting big data want as much data as possible, to use as the basis of the models businesses create to explain, predict, and affect behavior. In general, more data means better models. For this approach, the key operational mandate (all other things being equal) is to "gather it now, figure out why later." Fundamentally, this business approach is inconsistent with the idea behind most privacy rules.

In considering how these challenges apply to the health care industry, it is critical to recognize that:

- Much of the big data discussion is outside of the context of health care; BUT
- There is a wide variety of health care information (both HIPAA regulated and not) that is being scrutinized in the context of big data; and
- There is a growing range of big data activities being conducted by health care entities, again both in and out of HIPAA.

Expressed Concerns

With that background, how should the health care industry think about big data? Over the last several years, there has been important policy thinking about the impact of big data in general and the specific impact across the health care industry. For example, a White House Task Force issued a significant and thoughtful report in May 2014, called "Big Data: Seizing Opportunities, Preserving Values," available **here**. This report included a key finding – among a generally balanced discussion of the risks and benefits of big data analytics. According to this report, "A significant finding of this report is that big data analytics have the potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, employment, health, education, and the marketplace." In addition, this report made clear that "The privacy frameworks that currently cover information now used in health care may not be well suited to address these developments or facilitate the research that drives them."

The FTC sounded similar alarms, while also recognizing benefits to consumers and industry from big data. In its January 2016 report "Big Data: A Tool for Inclusion or exclusion," available **here**, the FTC concluded that "[t]he analysis of [big] data is often valuable to companies and to consumers, as it can guide the development of new products and services, predict the performance of individuals, help tailor services and opportunities and guide individualized marketing." In addition, "[a]t the same time, advocates, academics and others have raised concerns about whether certain uses of big data analytics may harm consumers, particularly low-income and underserved populations."

Potential Enforcement Sources

Companies thinking about these big data issues – or applying them without thinking about them – should consider some key elements.

First, much of the big data discussion for the health care industry occurs at the border of regulated and unregulated. The HIPAA rules may apply, but often will not.

However, there also are a variety of other laws and regulations that may be relevant, particularly the broad and general ability of the FTC to take action against "unfair and deceptive" practices, that certainly can include various approaches related to privacy and data analytics.

Second, while the core health care industry focuses on the U.S. Department of Health and Human Services (HHS) Office for Civil Rights (OCR) as a primary regulator, for big data activities over the next few years, it may be more important to think about the growing "enforcement" authority of two other groups – the State Attorneys General and the plaintiffs' bar. It is widely expected that the State AGs will play a more aggressive role in privacy and data security enforcement over the next few years. The New York Attorney General, for example, already has started to define practices for the "unregulated" health care industry. *See* Nahra, "New York Attorney General Addresses Key Health Care Privacy Gaps," *Privacy in Focus* (April 2017), available **here**.

Also, the plaintiffs' bar is gearing up to take action against inappropriate or suspicious big data practices. *See* Roberts, "These Popular Headphones Spy on Users, Lawsuit Says," *Fortune* (April 19, 2017), available **here**. While the question of "harm" will still be relevant to these actions, the plaintiffs' bar presents a real and ongoing risk for companies pushing the envelope on big data analytics.

Expanding Sources of Health Care Information

Part of the regulatory challenge for the health care industry stems from the limited scope of the HIPAA rules. HIPAA has always been a limited scope privacy/security rule - It applies to health care information only where a "covered entity" is involved. *See* Nahra, "Refresher on the HIPAA Privacy and Security Rules," *Privacy in Focus* (May 2017), available **here**. Accordingly, there always have been gaps where various entities collect or maintain health care data but are not covered by the HIPAA rules.

In the years since HIPAA went into effect, the amount of "health related" information that is being created and collected outside of the HIPAA rules has been growing, and growing in recent years in significant ways. For example, we now see:

- Web sites gather and distribute health care information without the involvement of a covered entity. These range from commercial web sites (e.g., Web MD) to patient support groups to sites that offer "personal health records" directly to consumers.
- These web site activities also have been supplemented by other web sites that are gathering and using information from a wide variety of vehicles to be used in relation to health care activities, including

significant and growing activities involving Google, Facebook, and Apple.

- We also have seen a significant expansion of mobile applications directed to health care data or offered in connection with health information.
- We have seen a broad variety of health care information gathered by "wearables" and other devices (whether regulated or not) that collect and create health care information.

At the same time that we have seen the growth of "non-HIPAA" health care information, we also are seeing an emerging (and related) issue – companies within the health care industry (and therefore regulated by HIPAA) bringing into their company information created outside of the health care system, much of it not obviously related to health care, but bringing it in to use for data analytics related to health care. Consider these recent headlines.

- "Your Doctor Knows You're Killing Yourself. The Data Brokers Told Her." "You may soon get a call from your doctor if you've let your gym membership lapse, made a habit of picking up candy bars at the check-out counter or begin shopping at plus-sized stores." *See* Pettypiece and Robertson, "Your Doctor Knows You're Killing Yourself. The Data Brokers Told Her," *Bloomberg Technology* (June 26, 2014), available here (subscription needed).
- "When a Health Plan Knows How You Shop." Health plan prediction models using consumer data from data brokers (e.g., income, marital status, number of cars), to predict emergency room use and urgent care. *See* Singer, "When a Health Plan Knows How You Shop," *The New York Times* (June 28, 2014), available **here**.
- "Bosses Tap Outside Firms to Predict Which Workers Might Get Sick." Employee wellness firms and insurers are working with companies to mine data about the prescription drugs workers use, how they shop, and even whether they vote, to predict their individual health needs and recommend treatments. *See* Silverman, "Bosses Tap Outside Firms to Predict Which Workers Might Get Sick," *The Wall Street Journal* (February 17, 2016), available **here** (subscription needed).

So, it is clear that the health care industry – and many others who use, collect, and analyze health care data or other data that can be used for health care purposes – have a variety of challenges to consider. First, what are the regulatory requirements? Second, where is the law moving in this area? And third, what should I be doing – considering ethics, customer relations, business strategy, and public relations – to address these issues over the next few years where there is a reduced likelihood of thoughtful and effective regulation of these topics?

On a broad level, through the White House Big Data report, the FTC's Data Broker report and otherwise, substantial concerns have been raised about how this data is being used, in contexts that raise questions about how health care services are provided and appropriate rights and protections for individuals in connection with their health care and their privacy. And, while this issue may be temporarily on hiatus, with the anti-regulation current Administration and Congress, this HIPAA/non-HIPAA issue is not going away, nor is the data analytics industry involving health care information slowing down in any way. There is too much data being used by too many people in too many risky contexts. In addition, there is growing pressure from a variety of audiences to "do something" about this non-HIPAA health care data.

Regulation Expansion Options

On the legislative/regulatory front, there seem to be three main options.

First, we could address the specific "gap" in the law – health related data that is not covered by HIPAA. This would involve a law and/or regulation creating privacy and security rules for "non-HIPAA" health care information. This would address this gap, and create rules for this information, but would leave two different sets of rules in place, for HIPAA and non-HIPAA data.

Second, we could pass a law or regulation addressing all health care information. This would place the use and disclosure rules for health care information on a single footing – avoiding issues at the margins, and creating both disparities in rights and differing commercial obligations depending on which side of the line you were on. The health care industry needs to pay close attention to this option – if a single "health care" law was passed, it might not look like HIPAA. If the health care industry likes the HIPAA rules, it needs to formulate a strategy to expand HIPAA to this broader environment – which isn't as easy as it sounds.

Third, and considering the fact that income, number of cars, marital status, and a variety of other factors are being determined to be valuable in gathering and analyzing health care insights and ongoing practices – maybe it is no longer possible to define "health care information" in a way that provides appropriate protections. This would lead to the broadest "solution" – a single privacy law/regulation covering all information. This "one size fits all" system is the preferred approach in other countries that have privacy laws (particularly the European Union), but has not been the approach in the United States so far, where we have preferred a "sector and practice" specific sets of laws and regulations. *See* Nahra, "Is the Sectoral Approach to Privacy Dead in the U.S.?" *Bloomberg BNA Privacy and Security Law Report* (April 4, 2016), available **here**.

While this debate about new laws and regulations continues (and perhaps stagnates for a few years), what else should companies be considering for the time being?

Data De-Identification

One key issue for consideration is the issue of "de-identification," including both an understanding of what this means and an analysis of whether appropriate de-identification techniques can avoid or reduce some of the risks associated with big data analytics.

While this is not the place for a full discussion of the HIPAA de-identification rules, the approach spelled out in HIPAA permits "protected health information" to be "de-identified" so that the information is no longer regulated by HIPAA, because it is no longer reasonably associated with an identifiable individual. Under HIPAA, a covered entity or other recipient of data can do anything with de-identified data. This "unrestricted" environment is one reason this information is so valuable – can be sold, used for research or for any other purpose under current law. This means that de-identified data can be used for:

- Research;
- Public Health;
- Big Data Analytics;
- Marketing; and
- Any other purpose if the data is appropriately de-identified.

Accordingly, the HIPAA rules – which require a strict standard for de-identification - present opportunities to engage in big data analytics without meaningful privacy risk. At the same time, one of the key challenges of big data analytics is addressing not only the "privacy risk," where an individual's identifiable data is used in ways that impact that individual, but also the "discrimination" concerns raised by the White House and others, where one person's data (whether de-identified or not) is used to create a model that is then applied to someone else. While these discrimination concerns are real, they are not typically "privacy" concerns, because the impacted individual often is not the one whose data was used to create the model.

Two Big Issues

In addition to these de-identification opportunities, companies in the health care industry and related industries – and their lawyers – will need to consider two other big picture issues:

Data Quality. One challenge for health care is that much of the data being brought into the health care system from other sources does not have the same level of quality as most health care information. Obviously, the importance of data quality depends on the purpose for which the data is being used. A bad marketing profile results in bad marketing. A bad clinical protocol kills someone. So it will be critical to evaluate data sources and data quality to the extent that this data is being used for meaningful health care purposes. Health care companies may find that better laws

dictating specific practices may lead to better quality data from many of these sources.

Ethics. While ethics may seem a quaint and antiquated notion in our current political environment, it is clear that ethics remains a critical principle within the health care industry. The vagaries and ambiguities of the legal structure today may place more importance on these ethics rather than less. Lawyers – in particular – will need to pay attention to these ethical issues, so that they can raise appropriate questions and concerns when businesses push for more and more aggressive use of data in a regulatory vacuum. Because of the broad risks of big data analytics – and the likely challenges from regulators, plaintiffs' attorneys, the media, and others – the fact that there are legal gaps and ambiguities cannot be an excuse for proceeding in an aggressive and unthinking way simply because there is no law that says you can't.

This question of big data for the health care industry is growing in importance every day. Businesses, academics, policymakers, and others are – every day – developing new thoughts and approaches to gathering, identifying, analyzing, and applying big data analytics to a broader variety of information. We don't know at this point what data will matter or how it will be used (and whether these conclusions will be accurate). Businesses – and their strategists and lawyers – need to be thinking about these big picture issues, and evaluating an approach that is consistent with both current law and likely future development of the law, but also with thoughtful business activities and emerging best practices related to privacy, data security, and overall implications of big data analytics.