WITTIE, LETSCHE & WALDO, LLP
915 15th Street, NW – 2nd Floor
Washington, DC 20005
Phone: 202-464-9350

February 15, 2011

VIA ONLINE SUBMISSION

Federal Trade Commission
Office of the Secretary
600 Pennsylvania Ave NW
Washington DC 20580

Re: A Preliminary Staff Report on "Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers"

We welcome the opportunity to submit these comments to the Federal Trade Commission on behalf of Wolters Kluwer Pharma Solutions.

Wolters Kluwer Pharma Solutions provides information and business intelligence for students, professionals and institutions in medicine, nursing, allied health and pharmacy, as well as offering marketing and publications services, business intelligence products, and advanced analytical tools and services to support life sciences professionals from drug discovery through development and distribution. Our customers include hospitals and libraries, research organizations, medical professionals, students, and pharmaceutical companies.

De-Identified Prescription Data Sets

Among our products are patient de-identified prescription data sets that we out-license to health care customers, including pharmaceutical companies, researchers, and government agencies. These data sets are composed of relevant information about prescriptions, including name of drug, dosage, and identifiable prescriber information, and they are always fully de-identified as to patients according to standards set out in Health Insurance Portability and Accountability Act (HIPAA) regulations.

These HIPAA de-identified prescription data sets are used by pharmaceutical and life science companies and researchers to inform basic medical research, speed drug development, and promote efficiencies in pharmaceutical marketing. They are also used by a wide array of public health and government agencies, including the Food and Drug Administration, the National

Institutes of Health, the Centers for Disease Prevention and Control, the Drug Enforcement Administration, the Centers for Medicare and Medicaid Services, and state and local health payers, public health, and law enforcement authorities to assist governmental initiatives ranging from tracking drug safety, conducting and regulating clinical research, reducing treatment disparities, detecting prescription abuse and financial fraud, and conducting quality and cost-effectiveness studies.

Although the government, research, public health, and nonprofit uses are vital for advancing societal goals, as explained further below, the prescription data sets themselves would not be commercially viable – and thus not available at all – if our commercial customers were not able to buy them and use them in commercial operations. Furthermore, there is no alternative, public sector source of this information, for developing HIPAA de-identified prescription data sets is an expensive and complex process. Commercially available data sets developed by healthcare information companies like Wolters Kluwer Pharma Solutions and others offer the only comprehensive sources of prescription information.

It is this valuable information product – HIPAA de-identified prescription data sets – that is relevant to the comments we submit to you today.

Scope of our Comments

We appreciate the leadership the Federal Trade Commission is showing by issuing the thoughtful staff report seeking to adapt our national privacy framework to a rapidly evolving consumer environment. We are limiting our remarks here to one narrow but vital issue addressed in the report – the scope of the proposed framework, which is proposed to cover all consumer data "that can be reasonably linked to a specific consumer, computer, or other device," as well as the related assertions in the report about the "decreasing relevance of the distinction between PII [Personally Identifiable Information] and non-PII."

We do not think that, rightly understood, health data that has been de-identified according to the formal HIPAA standard would be within the scope of the proposed framework, insofar as it would not be considered "reasonably linked to a specific consumer" (or computer or device.) Because of the vital importance, however, of de-identified health data streams to the national imperatives of improving health outcomes, finding new treatments, and increasing the cost-effectiveness of health care, we want to take this opportunity to emphasize how HIPAA de-identified data differs substantially from many purportedly anonymized consumer data sets. We would urge the Commission to resolve uncertainty within the health care environment by confirming that HIPAA de-identified health data should not be subjected to a new, additional regulatory structure, which could limit its availability for crucial societal uses.

The HIPAA De-Identification Standard

The HIPAA de-identification standard was carefully developed by numerous governmental experts and outside stakeholders over a number of years. The intent was to find the right way

to fully protect patient privacy while making health data still useful for a variety of important research and societal purposes.

Two methods of de-identifying under HIPAA. HIPAA set the standard for de-identified Protected Health Information (PHI) as "health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual."[1] HIPAA permits de-identification to be established two ways: (a) the expert statistician method and (b) the Safe Harbor method of accomplished prescribed identifiers.

The "expert statistician method" provides that PHI is de-identified only if "a person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not readily identifiable: (i) Applying such principles and methods, 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 (ii) Documents the methods and results of the analysis that justify such determination."[2]

The "Safe Harbor method" requires that the covered entity 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, and further requires that **all** of the following 18 identifiers be removed:

(1) Names;
(2) All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census:
    (a) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
    (b) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.
(3) All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older;
(4) Telephone numbers;
(5) Fax numbers;
(6) Electronic mail addresses;
(7) Social security numbers;
(8) Medical record numbers;

[1] 45 CFR §164,514(a),
[2] 45 CFR §164,514(b(1).

(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 (IP) 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, except that a code or other means of record identification may be used, provided it is not derived from or related to information about the individual, it is not otherwise capable of being translated so as to identify the individual, and it is not used or disclosed for any other purpose, and the mechanism for re-identification is not disclosed by the covered entity.[3]

Differences between HIPAA De-Identification and the Level of Anonymization Used in Recent Successful Re-Identification Attacks

Not all anonymized data sets are created equal. Nor is the risk of re-identification by an attacker equal for all anonymized data sets. A careful study of the requirements for de-identification per either of the two allowable HIPAA methods outlined above reveals that a high degree of diligence and care, and often expense and technological expertise, is required to convert data from PHI to de-identified.

We note that a number of FTC roundtable panelists asserted, and the report writers seem to agree, that the traditional distinctions between PII and non-PII have eroded and that information practices and laws that rely on the PII/non-PII distinction are "losing their relevance." There can be little doubt that such comments, which seem ubiquitous in privacy discussions today, have some validity insofar as they apply to the consumer information market today. Browser and device identifiers and machine fingerprinting technology, the aggregation of disparate, loosely anonymized consumer databases, expanding computing power, and the growing motivation to track and re-identify consumers (particularly given the financial pressures on free online content providers), and a few highly publicized, successful re-identification attacks are among the factors that lead some observers to conclude that the PII/non-PII distinction is now blurry and perhaps should not be recognized in legal regimens.

Nonetheless, we think these conclusions are less supportable when it comes to HIPAA de-identified data. The degree of statistical rigor and the extensive redaction of data elements required to meet the HIPAA standard is sometimes not well understood in the consumer privacy space, where there is no recognized standard at all for anonymization, let alone a statistically robust one. Further, the underlying details in the cases and anecdotes where data

[3] 45 CFR §164.514 (b)(2).

intended to be anonymous was re-identified can be distinguished from cases in which HIPAA de-identification is employed.

In analyzing the differences between HIPAA de-identification and unstandardized consumer data anonymization methods, several items should be noted:

(1) *Latanya Sweeney's work re-identifying certain health data*. Dr. Sweeney's work, which was cited in the staff report, in which she successfully re-identified a number of patients[4] is often erroneously cited as an example of the purported weaknesses of the HIPAA de-identification standard. Actually, the opposite is true. Prior to HIPAA's effective date, Dr. Sweeney obtained a copy of the Massachusetts Group Insurance Commission (GIC) database that contained highly detailed information about state employees' health. In fact, the data set contained *almost one hundred* attributes per patient encounter, <u>including the crucial attributes of zip code, gender, and date of birth</u>.[5] Because this data set was believed to be anonymous, the state GIC made it available for free to researchers and sold it to industry. For $20, Dr. Sweeney then purchased the voter registration rolls for Cambridge, Massachusetts, and succeeding in partially linking the data sets by searching for matching zip code, gender, and birth dates. Famously, among the names she successfully re-identified was then-Governor William Weld; Dr. Sweeney then mailed his medical records to his office.

The critical insight from this case is that the data Dr. Sweeney re-identified was <u>not</u> HIPAA de-identified, nor did it even approach the standard for HIPAA de-identification. To the contrary, her research results influenced the development of the HIPAA standard, resulting in the mandatory omission of zip code and date of birth attributes before data sets can qualify as de-identified under the Safe Harbor method.[6]

This case does raise an important question about the wisdom of releasing certain state health record databases "into the wild" without restrictions, given that these databases can serve as the crucial link for those attempting re-identification attacks. Dr. Sweeney reports that 37 states have legislative mandates to collect hospital level data and 17 states also collect ambulatory care data,[7] and states

[4] Latanya Sweeney, *Achieving k-Anonymity Privacy Protection Using Generalization and Suppression*, 10 Int'l. J. on Uncertainty, Fuzziness, and Knowledge-Based Systems, 571, 572 (2002).

[5] Statisticians have calculated that between 61% and 87% of the U.S. population can be uniquely identified by the combination of their zip code, date of birth, and gender. In 2000, Latanya Sweeney found the percentage was 87%, while Phillippe Golle's recalculations in 2006 found 61% for 1990 census data and 63% for 2000 data. Latanya Sweeney, *Uniqueness of Simple Demographics in the U.S. Population, Laboratory for International Data Privacy Working Paper*, LIDAP-WP4(2000); Phillippe Golle, *Revisiting the Uniqueness of Simple Demographics in the US Population*, 2006 Workshop on Privacy in the Elec. Soc'y Proc.

[6] Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. Rev. 1701 (2010).

[7] Sweeney, *supra*, note 3.

vary in the degree to which they publicly disclose the data they collect. While these state operations are beyond FTC jurisdiction, an analysis of actual re-identification risks should take a hard look at the appropriateness of states' making lightly anonymized health data widely available. Just as policymakers have learned to be more cautious about uploading drivers' license databases and court records containing Social Security numbers to the web, the public disclosure of lightly anonymized health data – especially data that would never approach the HIPAA de-identification standard – should be thoughtfully reviewed.

(2) *A study asserting the re-identifiability of certain genomic data.* Bradley Malin and Latanya Sweeney studied the susceptibility to re-identification of genomic data, either pseudonymous or believed to be anonymous, when released into a distributed healthcare environment under certain circumstances.[8] The salient fact relevant here is that the reference database Drs. Malin and Sweeney used to establish their data linkage was, again, a publicly available state hospital discharge system. The Illinois database, which included more than 99% of hospital discharges in the state from 1990 through 1997, included date of birth, gender, zip code, hospital visited, and clinical diagnosis and procedure codes. Obviously, this data would not have satisfied the HIPAA de-identification standard. In fact, if a hospital, instead of a state, had released patient data that included these identifiers, that could have constituted a grave HIPAA violation triggering substantial penalties. The point is that such a study reveals nothing about the privacy-protective efficacy of the HIPAA de-identification standard or the supposed blurring of lines between identifiable and HIPAA de-identified data, and it certainly should not lead one to any particular policy conclusions regarding HIPAA de-identified data.

(3) *The Netflix case.* In a well-publicized contest intending to use crowd-sourcing to improve the accuracy of Netflix's movie rating algorithms, the company released nearly 100 million movie ratings provided by almost 500,000 of its users, with direct identifiers of the users removed. Two researchers then published a paper demonstrating that it was possible to single out individual Netflix users (note, we did not say "identify") by linking them to publicly available ratings posted by registered users of the Internet Movie Database (IMBd).[9] Not surprisingly, the more obscure the movies ranked in both data sets, the easier it is to link the data sets. This is intuitive – if, for example, a movie was so extremely obscure that it had been ranked only *once* by a user in both databases, and the ranking had occurred on the same day, the statistical adversary would have achieved a very high probability of a successful linkage. In any case, what the linkage would have

[8] Bradley Malin and Latanya Sweeney, *How (not) to protect genomic data privacy in a distributed network: using trail re-identification to evaluate and design anonymity protection systems*, 37 J. of Biomedical Informatics 179 (2004).

[9] Arvind Narayanan and Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets*, 2008 IEEE Symp. on Security and Privacy 111.

revealed is the publicly available username registered on the IMDb. The paper gives no hint that further attacks somehow linked that IMDb username to an actual name unless the user chose to register in her real name – or how such an attack linking usernames and actual identities could be done.

There are, of course, a number of conclusions one may draw from the Netflix case, including: the vulnerability to re-identification attacks of highly granular data sets containing what the authors call *micro-data* (information about specific people, not aggregated statistics); the prudence a dataholder should take before releasing such data sets; the appropriateness of releasing such data sets only to carefully selected recipients rather than publicly; whether release of such datasets should always be barred by contractual bans on, and penalties for, re-identification attacks; and the caution a consumer should exercise before posting granular and/or obscure information about oneself, especially under her real name. What does <u>not</u> follow, however, is any conclusion that the Netflix case somehow disproves the protectiveness of the HIPAA de-identification standard. Consider how a rigorous de-identification standard analogous to that of the HIPAA expert statistician method could apply to consumer data such as movie ratings. Under HIPAA, an expert statistician must analyze the risk that information "could be used, alone <u>or in combination with other reasonably available information</u>," to re-identify individuals. The existence of the highly detailed, publicly available movie ranking database, now buttressed by the paper's revelation of the attack methodology, would presumably make the re-identification risk too high to qualify under a standard as rigorous as that in HIPAA.

(4) *Absence of evidence of <u>any</u> successful re-identification attacks on HIPAA de-identified data*. We are unaware of any instance reported in a professional journal in which data that meets the HIPAA de-identification standard has been successfully re-identified. In fact, on January 4, 2010, the Department of Health and Human Services solicited bids from statisticians or others to demonstrate methods and technologies to re-identify a de-identified data set.[10] The results of that challenge have not yet been announced. The existence of any threat to patient privacy from HIPAA de-identified data is thus theoretical at this time, and any potential vulnerability of HIPAA de-identified data should be analyzed separately from the re-identification attacks discussed above. But even if a small risk were to be proven in the future, policymakers should carefully weigh such a risk against the wide array of important benefits of de-identified health data to patients, the economy, and taxpayers.

[10] US Department of Health and Human Services. Comprehensive research on re-identifying a HIPAA de-identified dataset [Internet].Washington (DC): HHS; 2010 Jan 4 [cited 2010Aug 12]. Available from: https://www.fbo.gov/index?s=opportunity&mode=form&id=bf5b42d4d605295ec2d4bde88078cffa&tab=core&_cview=0&cck=1&au=&ck

Use of HIPAA De-Identified Data Enhances Privacy and Should Be Encouraged

Both common sense and recent computer science research support the idea that patient privacy is most at risk when data is fully identifiable and that that risk decreases as data becomes more and more strictly anonymized. We thus agree with the observation by the Center for Democracy and Technology (CDT) that "[u]se of the least identifiable data should always be encouraged . . . ."[11] Safeguarding patient data by removing identifiers where not needed for the purpose to be served makes sense. In some cases, the research utility of data sets is severely hampered if the data is scrubbed sufficiently to meet the strict HIPAA de-identification standard. In anticipation of this problem, the drafters of HIPAA created a less strict version of anonymization, the Limited Data Set, which allows the inclusion of two more identifiers than the de-identification Safe Harbor permits. Using Limited Data Sets involves legally mandatory restrictions on uses and certain contractual controls, including a contractual ban on re-identification.[12] In the case of the prescription datasets that we provide, however, meeting the tighter de-identification standard is not problematic, for our customers have no need for, and are not interested in, patient identities. Because using fully de-identified data for a particular purpose will result in far better privacy protections for individuals than using more identifiable or fully identifiable data, our public policy should protect consumers by encouraging use of *more* HIPAA de-identifiable data, not imposing new impediments on it.

De-identification is an important compliance tool for HIPAA Covered Entities and Business Associates. The Health Information Technology for Economic and Clinical Health (HITECH) Act, passed as part of the American Recovery and Reinvestment Act of 2009, significantly expanded the already existing incentives for health entities to fully de-identify their data. The penalties for impermissible uses and disclosures of PHI (*i.e.*, non-de-identified health data) were substantially increased, with both Attorneys General and HHS now authorized to bring enforcement actions. The new data breach laws also incentivize health entities to de-identify data wherever feasible, for the accidental release of de-identified data does not trigger expensive data breach notice obligations. Because robust de-identification safeguards the privacy of patients and consumers, these trends should be encouraged. As stated by HHS in its original commentary setting forth the HIPAA de-identification standard:

> Indeed, it would be our hope that covered entities, their business partners, and others would make greater use of de-identified health information than they do today, when it is sufficient for the research purpose. Such practice would reduce the confidentiality concerns that result from the use of individually identifiable health information for some of these purposes.[13]

The Societal Value of HIPAA De-Identified Data

[11] *Encouraging the Use of, and Rethinking Protections for De-Identified (and "Anonymized") Health Data,* Center for Democracy and Technology, June 2009, available at http://www.cdt.org/healthprivacy/20090625_deidentify.pdf.
[12] 45 CFR §164.514(e).
[13] 64 Fed. Reg. at 59946.

Although not well understood by the public, the investments made by the health information industry in gathering, managing, and de-identifying health data have served as the foundation of considerable health care system improvements in recent years. Specifically, HIPAA de-identified prescription data sets drive activities such as:

- Quality improvements – Researchers use de-identified prescription data to assess treatment outcomes, establish standards of care, and improve evidence-based medicine;
- Research – Clinical researchers use de-identified prescription data to design clinical trials and to identify providers who are treating possible trial participants; population-level epidemiological researchers uses the data to find statistical associations and predictive patterns to identify disparities and improve care;
- Public health – De-identified prescription data is used for adverse drug event monitoring, syndromic surveillance, and early detection of infectious outbreaks or even bioterrorism incidents;
- Cost-effectiveness improvements – De-identified prescription data is used in comparative effectiveness and other efficiency studies, which have strong support in the current Administration and are crucial for bending the health care cost curve downward; and
- Commercial uses – many companies use de-identified prescription data to develop and improve their products and support their core operations, such as pharmaceutical companies using the databases to learn which aggregate populations are using specific drugs, analyze safety risks, and improve the efficiency of sales efforts.

When the HIPAA de-identification standard was being established, HHS understood that de-identified data would fuel health care improvements, as expressed in their commentary:

> There are many instances in which such individually identifiable health information is stripped of the information that could identify individual subjects and is used for analytical, statistical and other related purposes. Large data sets of de-identified information can be used for innumerable purposes that are vital to improving the efficiency and effectiveness of health care delivery, such as epidemiological studies, comparisons of cost, quality or specific outcomes across providers or payers, studies of incidence or prevalence of disease across populations, areas or time, and studies of access to care or differing use patterns across populations, areas or time. Researchers and others often obtain large data sets with de-identified information from providers and payers (including from public payers) to engage in these types of studies. This information is valuable for public health activities (*e.g.*, to identify cost-effective interventions for a particular disease) as well as for commercial purposes (*e.g.*, to identify areas for marketing new health care services).[14]

[14] *Id.*

Controversies and Open Questions Involving HIPAA De-Identified Data

In recognition of the evolving consumer data environment and the publicity about re-identification cases, Congress tasked HHS in HITECH with issuing new guidance on how best to implement the requirements for the Safe Harbor and expert statistician methods of de-identifying PHI. As part of that effort, HHS held a two-day workshop in March, 2010 to hear from statisticians and biomedical, policy, and legal experts regarding de-identification techniques and best practices, re-identification risks, and potential policy changes. While much attention was focused on the academic research involved in published re-identification attacks, there was also discussion about how to balance the privacy risks potentially caused by insufficiently de-identifiable data with the risks to patients and society from poor treatment, research, and information resulting from new barriers to de-identified health data. Concern was expressed that a single-minded focus on theoretical privacy risks – thus far shown only in academic publications and not even involving HIPAA de-identified data – could grow into serious impediments to the availability of crucial information needed to achieve breakthrough treatments and speed health system improvements.[15]

A number of biomedical researchers at the HHS de-identification workshop expressed their view that the current de-identification standard already inappropriately impedes research, and that any new impediments or restrictions should be opposed as harmful to patients' interests. Concerns cited included the difficulty of conducting peer reviews of research due to blocked access to the data involved in the original study, the near-impossibility of obtaining consent for retrospective review of records where de-identified data would not achieve the research protocol objectives, and the obstacles posed to linking records needed for personalized medicine research. These concerns echo ones frequently expressed by the biomedical research community. For example, an Institute of Medicine panel recently concluded that HIPAA "as currently implemented . . . impedes important health research."[16]

Steps to enhance patient privacy without disrupting and burdening research are under discussion, both within HHS and in policy circles. Last year the Confidentiality Coalition of the Healthcare Leadership Council, in which Wolters Kluwer Pharma Solutions participates, wrote to HHS urging that the agency recommend several best practices, including (a) use of contract

[15] The issue of HIPAA de-identified prescription data is also involved in a case currently before the Supreme Court. The case of *IMS Health et al v. Sorrell* involves the constitutionality of state laws restricting commercial uses of prescriber-identifiable, but patient de-identified, data. The lower courts have agreed that the use of this data does not raise patient privacy issues: for example, the First Circuit Court of Appeals decision upholding one of the state laws stated that "the regulation does not in any cognizable way touch on the privacy of the examination room;" The U.S. District Court in Maine said there was 'no evidence that the current practices of the [companies] have had or realistically could have any effect on patient confidentiality;" and the U.S. District Court in New Hampshire pointed out that the state Attorney General "does not claim that data is being exploited to compromise patient privacy." Instead, the legal questions at issue focus on the First Amendment.

[16] Sharyl J. Nass, Laura A. Levit, and Lawrence O. Gostin, Editors, Committee on Health Research and the Privacy of Health Information, Institute of Medicine, *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*, National Academies Press, 2009.

language banning recipients of de-identified datasets from attempting re-identification of individuals; and (b) reasonable and appropriate safeguards for the protection of de-identified data.[17] The Coalition also expressed concern about privacy vulnerabilities arising from possible misuse of publicly available reference databases such as state hospital discharge databases. A recent Health Affairs article recommended a new privacy framework for information-based health research (as opposed to clinical trials) by creating a research safe harbor in which certain legal obstacles would be lessened in exchange for heightened safeguards and outside audits.[18] CDT has issued a paper supporting the use of de-identified health data and recommending several steps that could improve privacy protections.[19] CDT's recommendations include re-examining the HIPAA standard, strengthening accountability through data use agreements over de-identified data, incentivizing the use of less-than-fully-identifiable data, and encouraging the use of limited access datasets and other technical solutions.

At Wolters Kluwer Pharma Solutions, we support and, in fact, already implement many of these safeguards on the de-identified prescription databases we maintain. We ensure that the databases are fully de-identified per the HIPAA standard, we protect them with strong security controls, and we license them subject to contractual controls.

Recommendation

On a policy level, we support the focused attention HHS and numerous expert stakeholders are applying to the question of ensuring that de-identified health data is both readily available for a myriad of societally beneficial purposes and is well protected from a privacy standpoint. In general, we think the existing HIPAA de-identification standard reflects an appropriate balancing of individual and societal interests. That said, to the extent that there is evidence of new problems or risks arising, we welcome attention to modifications in de-identification policy, such as contractual controls, security safeguards, and legal penalties for malevolent re-identification attacks. What would concern us, however, is the idea of subjecting de-identified health data to a new and divergent regulatory framework. It is clear that, for at least a decade, an enormous amount of scarce health care resources have gone into developing the regulatory framework and health system infrastructure for HIPAA de-identification, that HIPAA de-identification is a vital compliance tool for hospitals, providers, and insurers, and that HIPAA de-identification does not threaten but rather enhances patient privacy. Most of all, patients and society as a whole need de-identified health data, including prescription data, to be widely available to appropriate users and, in fact, to be more extensively used in the future to improve the quality, safety, and cost-effectiveness of health care.

---

[17] Healthcare Leadership Council Confidentiality Coalition. *Principles on de-identification and use/disclosure of de-identified data sets* Mar 2009. Available from: http://www.acrohealth.org/acro-leads-coalitionon-data-use-in-research-hitprivacy.html

[18] Douglas Peddicord, Ann B. Waldo, Marc Boutin, Tina Grande, and Luis Gutierrez Jr., *A Proposal to Protect Privacy of Health Information While Accelerating Comparative Effectiveness Research*, Health Affairs 29:11, 2082-90 (2010).

[19] Center for Democracy and Technology, *supra,* note 11.

For these reasons, we would encourage the Commission to consider:

- Carefully distinguishing HIPAA de-identified data from any broad assertions or conclusions about the privacy vulnerability of non-PII, given that consumer data "anonymization," unlike HIPAA de-identification, is not subject to any rigorous standard or expert statistician review; and

- Expressly excluding both HIPAA-regulated Personal Health Information and HIPAA de-identified health data from the purview of any new consumer protection framework, in light of the extensive work already underway to evaluate whether new restrictions or best practices regarding HIPAA de-identified health data are appropriate and the obligation of HHS to issue new de-identification guidance.

Thank you for the opportunity to provide these comments, and we would be happy to follow up regarding any questions. We appreciate the thoughtful work the Commission is undertaking to update the national privacy framework in light of evolving consumer protection needs.

Yours very truly,

Ann B. Waldo
Wittie, Letsche & Waldo, LLP
915 Fifteenth Street NW, Second Floor
Washington, DC 20005
202-464-9350
awaldo@wlw-lawfirm.com

On behalf of Wolters Kluwer Pharma Solutions