

No. 10-779

---

---

IN THE  
*Supreme Court of the United States*

---

WILLIAM H. SORRELL,  
ATTORNEY GENERAL OF VERMONT, et al.,  
*Petitioners,*

v.

IMS HEALTH INC., et al.,  
*Respondents.*

---

On a Writ of Certiorari to  
The United States Court of Appeals  
for the Second Circuit

---

**BRIEF OF *AMICI CURIAE* ELECTRONIC  
PRIVACY INFORMATION CENTER (EPIC)  
AND LEGAL SCHOLARS AND TECHNICAL  
EXPERTS IN SUPPORT OF THE  
PETITIONERS**

---

MARC ROTENBERG  
*Counsel of Record*  
JOHN VERDI  
SHARON GOOTT NISSIM  
THOMAS H. MOORE  
ELECTRONIC PRIVACY  
INFORMATION CENTER (EPIC)  
1718 Connecticut Ave. NW  
Suite 200  
Washington, DC 20009  
(202) 483-1140

March 1, 2011

---

---

TABLE OF CONTENTS

TABLE OF CONTENTS ..... i

TABLE OF AUTHORITIES ..... iii

INTEREST OF THE *AMICI CURIAE* ..... 1

SUMMARY OF THE ARGUMENT ..... 9

ARGUMENT ..... 10

    I. Medical Privacy is a Fundamental  
        Concern for Patients. .... 10

    II. The Vermont Law Seeks to Safeguard  
        Medical Privacy, a Fundamental Concern  
        for those Receiving Prescription Drugs..... 13

    III. Data-Mining Companies Fail to Mitigate  
        Medical Privacy Risks to Vermont  
        Residents..... 19

        A. *The Cryptographic Technique Used to  
            Conceal the Identity of Patients is  
            Inadequate* ..... 21

        B. *Patient Records are At Risk of Being  
            Reidentified*..... 24

        C. *Other Data Is At Risk of Being  
            Reidentified*..... 29

        D. *Overturning the Vermont Statute Poses  
            a Risk of Increasingly Widespread  
            Access to Health Data*..... 33

    IV. Data-mining of Prescriber Information  
        Implicates the Constitutional Right to  
        Informational Privacy ..... 36

**CONCLUSION ..... 39**

## TABLE OF AUTHORITIES

### CASES

<i>Bartnicki v. Vopper</i> , 532 U.S. 514 (2001) .....	17
<i>Central Hudson Gas &amp; Elec. Corp. v. Pub. Serv. Comm’n of N.Y.</i> , 447 U.S. 557 (1980).....	15
<i>IMS Health Inc. v. Ayotte</i> , 490 F. Supp. 2d 163 (D.N.H. 2007) .....	10
<i>IMS Health Inc. v. Sorrell</i> , 630 F.3d 263 (2d Cir. 2010) .....	2, 27
<i>IMS Health v. Ayotte</i> , 550 F.3d 42 (1st Cir. 2008) .....	15
<i>NASA v. Nelson</i> , 131 S. Ct. 746 (2011).....	36, 37, 38
<i>Nat’l Cable &amp; Telecomms. Ass’n v. FCC</i> , 555 F.3d 996 (D.C. Cir. 2009) .....	17
<i>Northwestern Memorial Hospital v. Ashcroft</i> , 362 F.3d 923 (7th Cir. 2004).....	19
<i>Whalen v. Roe</i> , 429 U.S. 589 (1977).....	36, 37
<i>Wooley v. Maynard</i> , 430 U.S. 705 (1977) .....	17

### STATUTES

18 Vt. Stat. Ann. tit. 18, § 4631 (2010) .....	15
21 U.S.C. § 822 (2010) .....	14
Pub. L. No. 104-191 (1996).....	14
Vt. Acts & Resolves 80 (S.115) § 1(29).....	13
Vt. Stat. Ann. tit. 18, § 4631 (2010) .....	11, 13

**OTHER AUTHORITIES**

- Alessandro Acquisti and Ralph Gross,  
*Predicting Social Security Numbers from  
 Public Data*, 106 PROCEEDINGS OF THE NAT'L  
 ACAD. OF SCI., no. 27, 10975 (Jul. 7, 2009) ..... 32
- Anita Allen, *Privacy and Medicine*, Stanford  
 Encyclopedia of Philosophy (2009) ..... 16, 17, 18
- AOL Advertising, *Case Study: OTC Pharma  
 Leader Drives Offline Sales with AOL's  
 Online Targeting* ..... 34
- Arvind Narayanan and Vitaly Shmatikov,  
*Myths and Fallacies of "Personally  
 Identifiable Information,"* 53 COMM'N OF THE  
 ACM, 24-26 (June 2010) ..... 27
- Arvind Narayanan and Vitaly Shmatikov,  
*Robust De-anonymization of Large Datasets  
 (How to Break the Anonymization of the  
 Netflix Prize Dataset)*, Feb. 5, 2008 ..... 33
- Bert den Boer and Antoon Bosselaers, *Collisions  
 for the Compression Function of MD5*,  
 Proceedings of Eurocrypt '92, Advances in  
 Cryptology, 71-88 (1992) ..... 23
- BEYOND THE HIPAA PRIVACY RULE: ENHANCING  
 PRIVACY, IMPROVING HEALTH THROUGH  
 RESEARCH, 265 (Sharyl J. Nass et al, ed.,  
 2009) ..... 28
- BRUCE SCHNEIER, APPLIED CRYPTOGRAPHY 436  
 (2nd ed. 1996) ..... 22
- Bruce Schneier, *Forging SSL Certificates*,  
 Schneier on Security (Dec. 31, 2008) ..... 24

Chad Dougherty, <i>Vulnerability Note</i> <i>VU#836068: MD5 Vulnerable to Collision</i> <i>Attacks</i> , United States Computer Emergency Readiness Team (Dec. 31, 2008).....	24
Convention on Psychotropic Substances, art. 16, opened for signature Feb. 21, 1971, 1019 U.N.T.S 175 .....	14
David Chaum, <i>Achieving Electronic Privacy</i> , SCI. AM., 96-101 (1992) .....	28
Demand Media, Inc., “Form S-1,” August 6, 2010, p104 .....	34
<i>Discussion Draft of Health Information</i> <i>Technology and Privacy Legislation: Hearing</i> <i>on H.R. 6357 Before the Subcomm. on Health</i> <i>of the H. Comm. on Energy and Commerce</i> , 110 <sup>th</sup> Cong. (2008) (Testimony of Dr. Deborah Peel) .....	16
Fida Kamal Dankar & Khaled El Emam, <i>A</i> <i>Method for Evaluating Marketer Re-</i> <i>identification risk</i> , PROCEEDINGS OF THE 2010 EDBT/ICDT WORKSHOPS, ACM, Article 28 (2010) .....	25
Grigorios Loukides, et. al., <i>Anonymization of</i> <i>Electronic Medical Records for Validating</i> <i>Genome-wide Association Studies</i> , 107 PROCEEDINGS OF THE NAT'L ACAD. OF SCI., no. 17, 7898, (March 11, 2010) .....	30
Harris Interactive, <i>HIPAA Notices Have</i> <i>Improved Public’s Confidence That Their</i> <i>Medical Information is Being Handled</i> <i>Properly: However public split on benefits of</i> <i>and privacy risks associated with Electronic</i> <i>Medical Records (EMR)</i> , Feb. 24, 2005 .....	11

Homer N, et al., <i>Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays</i> . 4 PLOS GENETICS, no.8 (August 2008) .....	30, 31
Jerry Kang, <i>Cyberspace Privacy</i> , 50 STAN. L. REV. 1193, 1209 (1998) .....	18
Joe Mullin, <i>States Consider Limits on Medical Data-mining</i> , BOSTON GLOBE, Apr. 7, 2007 ....	11, 12
Khaled El Emam et al., <i>Evaluating Common De-identification Heuristics for Personal Health Information</i> , 8 J. MED. INTERNET RES. 4 (2006) .....	25, 26
Latanya Sweeney, <i>K-anonymity: A Model for Protecting Privacy</i> , 10 INTERNATIONAL JOURNAL ON UNCERTAINTY, FUZZINESS AND KNOWLEDGE-BASED SYSTEMS, no. 5, 557-70 (2000) .....	26
Latanya Sweeney, <i>Patient Identifiability in Pharmaceutical Marketing Data</i> , Cambridge, Data Privacy Working Paper No. 1015 (2011) .....	27
Latanya Sweeney, <i>Simple Demographics Often Identify People Uniquely</i> (Carnegie Mellon University, Data Privacy Working Paper No. 3, 2000) .....	24, 25
Latanya Sweeney, <i>Weaving Technology and Policy Together to Maintain Confidentiality</i> , 25 J. LAW, MED., & ETHICS 98, 98-99 (1997) .....	13
Mark Rothstein, <i>The Hippocratic Bargain and Health Information Technology</i> , 38 J. LAW, MED., & ETHICS 11 (2010).....	17, 18
Marketing Technology Solutions Corp. (MTS), <i>QH Connect</i> .....	34

Matthew D. Mailman, et al., <i>The NCBI dbGaP Database of Genotypes and Phenotypes</i> , 39 NATURE GENETICS, 1181 (2007) .....	30
Michael Barbaro & Tom Zeller Jr., <i>A Face is Exposed for AOL Searcher No. 4417749</i> , N.Y. TIMES, Aug. 9, 2006, at A1.....	32
Nat'l Physician's Alliance, <i>Issue Brief: The Sale of Physician Prescribing Data Raises Health Care Costs</i> (Feb. 2009) .....	12
<i>Patient Privacy Rights/Zogby International Poll</i> , Nov. 23, 2010 .....	16
Peter Wayner, <i>DISAPPEARING CRYPTOGRAPHY 32</i> (2nd ed. 2002) .....	23
Rachel Ehrenberg, <i>Hiding Patients in Plain Sight</i> , Science News, Apr. 12, 2010.....	30, 31
Razorfish LLC, <i>Outlook Report 2010</i> .....	34
Ron Rivest, <i>[Python-Dev] hashlib - faster md5/sha, adds sha256/512 support</i> (Dec. 16, 2005) .....	23
Ron Rivest, <i>The MD5 Message-Digest Algorithm</i> , RFC 1321 (Apr. 1992) .....	22
Single Convention on Narcotic Drugs, art. 19, 20, Mar. 30, 1961, 18 U.S.T. 1407, 520 U.N.T.S. 151 .....	14
Stolberg & Gerth, <i>High-Tech Stealth Being Used to Sway Doctor Prescriptions</i> , N.Y. TIMES, Nov. 16, 2000.....	14
Tanya Alberts, <i>Doctors Ask AMA to Assure Some Privacy for their Prescription Pads</i> , AMNews, Dec. 25, 2000 .....	12
The Prescription Project, <i>Prescription Data Mining Fact Sheet</i> , Nov. 19, 2008 .....	11



Vermont Medical Society Resolution, “Ensuring  
the Privacy of Prescription Information,” Oct.  
14, 2006..... 13

Vlastimil Klima, *Finding MD5 Collisions – A  
Toy For a Notebook* (Mar. 5, 2005)..... 20

**RULES**

45 C.F.R. § 164.312(e)(2)(ii) (2010) ..... 21, 33

45 C.F.R. § 164.502(d)(2) (2010)..... 13, 26

45 C.F.R. § 164.514(b)(2)(i)(B)-(C) (2010) ..... 25

45 C.F.R. §§ 164.312(e)(2)(ii) (2010) ..... 13

Vt. Bd. Of Pharmacy Admin. Rule § 9.1 ..... 13

**INTEREST OF THE *AMICI CURIAE*<sup>1</sup>**

The Electronic Privacy Information Center (“EPIC”) is a public interest research center in Washington, D.C. EPIC was established in 1994 to focus public attention on emerging civil liberties issues and to protect privacy, the First Amendment, and other Constitutional values.<sup>2</sup> The EPIC Advisory Board includes leading technical experts and legal scholars whose work has contributed to many of the techniques and policies that help safeguard privacy in the modern era.

For this reason, EPIC has participated as *amicus curiae* in many cases that concern emerging privacy issues before this Court, including *NASA v. Nelson*, 131 S. Ct. 746 (2011); *Doe v. Reed*, 130 S. Ct. 2811 (2010); *Quon v. City of Ontario*, 130 S. Ct. 2619 (2010); *Flores-Figueroa v. United States*, 129 S. Ct. 1886 (2009); *Herring v. United States*, 555 U.S. 135 (2009); *Crawford v. Marion County Election Board*,

---

<sup>1</sup> Letters of consent have not been lodged with the Court because on January 21, 2011, Respondents lodged with the Court their “consent to the filing of amicus curiae briefs, in support of either party or of neither party,” and on January 24, 2011, Petitioners lodged with the Court their “consent to the filing of amicus curiae briefs, in support of either party or of neither party.” In accordance with Rule 37.6, the undersigned states that no monetary contributions were made for the preparation or submission of this brief, and this brief was not authored, in whole or in part, by counsel for a party.

<sup>2</sup> EPIC Appellate Advocacy Fellow Conor Kennedy contributed to the preparation of this brief.

553 U.S. Ct. 181 (2008); *Hiibel v. Sixth Judicial Circuit of Nevada*, 542 U.S. 177 (2004); *Doe v. Chao*, 540 U.S. 614 (2003); *Smith v. Doe*, 538 U.S. 84 (2003); *Department of Justice v. City of Chicago*, 537 U.S. 1229 (2003); *Watchtower Bible and Tract Society of N.Y., Inc. v. Village of Stratton*, 536 U.S. 150 (2002); *Reno v. Condon*, 528 U.S. 141 (2000).

Concerning the matter before this Court, EPIC filed an amicus brief in the proceeding below, *IMS Health Inc. v. Sorrell*, 09-1913-CV L, 2010 WL 4723183 n.4 (2d Cir. Nov. 23, 2010) *cert. granted*, 131 S. Ct. 857 (U.S. 2011). Citing the EPIC amicus brief below, Judge Debra Ann Livingston stated in dissent, “[I]n an era of increasing and well founded concern about medical privacy and the rampant dissemination of confidential information, the federal government has repeatedly acted on that interest and legislated to protect the privacy of medical records.” *IMS Health Inc. v. Sorrell*, 630 F.3d 263 (2d Cir. 2010) (Livingston, J., dissenting), *cert. granted*, 131 S. Ct. 857 (U.S. 2011) (No. 10-779) (citations omitted). EPIC also filed an amicus brief in a similar case in the First Circuit, *IMS Health v. Ayotte*, 550 F.3d 42 (1st Cir. 2008) *cert. denied*, 129 S. Ct. 2864 (2009).

EPIC is joined in this brief by consumer, practitioner, and privacy organizations that have substantial expertise in issues concerning medical privacy and also support the privacy interests that the Vermont statute under consideration by the Court seeks to advance.

At issue in this case are the privacy interests of Vermont residents that the state of Vermont sought to protect through the enactment of legislation. The state has a vital interest in regulating conduct that

enables the transfer and sale of personal medical record information. Respondent IMS Health challenged this regulation as a violation of its asserted right to profit from the sale of this sensitive data. Amicus therefore submits this brief to make clear the substantial interest in safeguarding sensitive personal information as well as the related concern about the transfer of “anonymized” patient data to data-mining firms.

EPIC supports the outcome reached by the district court. In fact, EPIC believes that the court did not go far enough in stating the extent of the privacy interest at issue in this statute now under consideration by the Court. It is the nature of rapid technological change that the risks to personal privacy are often greater than can be readily understood at the time they emerge. This brief of *amicus* EPIC shows that in addition to the concerns expressed about the privacy of prescriber data, there are also substantial concerns for the privacy of patient data. Further, the techniques for anonymity adopted by Respondents do not adequately safeguard these interests. For these reasons, EPIC urges reversal of the appellate court’s decision and respectfully asks the Court to remand the case to the District Court so that the substantial privacy interests at issue in the transfer of medical information will be given sufficient weight.

*Technical Experts and Legal Scholars*

Dr. Alessandro Acquisti, Associate Professor of  
Information Technology and Public Policy,  
Carnegie Mellon University

Steven Aftergood, Senior Research Analyst,  
Federation of American Scientists

Grayson Barber, Esq., Grayson Barber, LLC

Francesca Bignami, Professor, George  
Washington University School of Law

Christine L. Borgman, Professor & Presidential  
Chair in Information Studies, UCLA

Stefan Brands, Adjunct Professor at McGill  
University School Of Computer Science

Dr. Whitfield Diffie, Dr. sc. techn. (hc), ScD (hc)

David Farber, Professor of Computer Science and  
Public Policy, Carnegie Mellon University

Addison Fischer, Former owner, RSA Data  
Security, Co-founder, Verisign

David H. Flaherty, Professor Emeritus of History  
and Law, University of Western Ontario;  
Information and Privacy Commissioner for  
British Columbia, 1993-99

Philip Friedman, Friedman Law Offices, PLLC

Deborah Hurley, Chair, EPIC Board of Directors

Ian Kerr, Associate Professor, Canada Chair of  
Ethics, Law, and Technology, University of  
Ottawa

Jerry Kang, Professor of Law, UCLA School of Law

Chris Larsen, CEO and Co-Founder, Prosper Marketplace, Inc.

Rebecca MacKinnon, Schwartz Senior Fellow, New America Foundation

Mary Minow, Library Law Consultant

Pablo Molina, Associate VP of IT and Campus CIO, Georgetown University

Helen Nissenbaum, Professor, Media, Culture & Communication, NYU

Ray Ozzie, (former) Chief Software Architect, Microsoft

Deborah C. Peel, MD, Founder and Chair, Patient Privacy Rights

Chip Pitts, Lecturer, Stanford Law School and Oxford University

Ronald L. Rivest, Professor of Electrical Engineering and Computer Science, MIT

Bruce Schneier, Security Technologist; Author, Schneier on Security (2008)

Latanya Sweeney, Visiting Professor, Harvard University, Distinguished Career Professor of

Computer Science, Technology and Policy,  
Carnegie Mellon University

Frank M. Tuerkheimer Professor of Law  
University of Wisconsin Law School

Edward G. Viltz, [www.InternetCC.org](http://www.InternetCC.org)

(Affiliations are for identification only)

*Privacy, Practitioner and Consumer Organizations*

American Psychoanalytic Association

The American Psychoanalytic Association is an association of 3,400 psychiatrists and psychoanalysts based in New York City whose members have practices throughout the United States, including Vermont.

Bill of Rights Defense Committee

The Bill of Rights Defense Committee is a national non-profit grassroots organization. We defend the rule of law and rights and liberties challenged by overbroad national security and counter-terrorism policies.

Center for Digital Democracy

The Center for Digital Democracy is a leading U.S. digital privacy non-profit organization that educates the public about the role of consumer data collection used for interactive

advertising, especially in the field of online health information and services.

#### Center for Media and Democracy

The Center for Media and Democracy is a national independent, non-profit, non-partisan media, policy, and consumer watchdog group located in Madison, Wisconsin.

#### Consumer Federation of America

The Consumer Federation of America is an association of non-profit consumer organizations that was established in 1968 to advance the consumer interest through research, advocacy, and education.

#### Patient Privacy Rights

Patient Privacy Rights is the nation's health privacy watchdog, a 501(c)3 nonprofit organization.

#### Privacy Activism

PrivacyActivism is a non-profit organization whose goal is to enable people to make well-informed decisions about the importance of privacy on both a personal and societal level.

#### World Privacy Forum

The World Privacy Forum is a nonprofit, non-partisan 501 (C) (3) public interest research group. The organization is focused on



conducting in-depth research, analysis, and consumer education in the area of privacy.

#### Privacy Rights Clearinghouse

Privacy Rights Clearinghouse is a nonprofit consumer organization with a two-part mission – consumer information and consumer advocacy.

## SUMMARY OF THE ARGUMENT

The Vermont Confidentiality of Prescription Information Law seeks to safeguard the privacy of prescribing information. This is a substantial state interest that is of even greater concern than the record below reveals. The “deidentification” technique adopted by the Respondents in this matter does not adequately safeguard the medical privacy of Vermont residents or the residents of other states whose personal prescribing information could be obtained by data-mining firms and subsequently sold to pharmaceutical companies. These records include the prescriber’s name and address; the name, dosage, and quantity of the drug prescribed; the date and location at which the prescription was filled; and the patient’s age and gender. The only missing element – the patient’s actual name – is concealed by a weak cryptographic technique that does not actually prevent reidentification of the patient by Respondent. In such circumstance, the Vermont law, and the many other similar state confidentiality laws, seek to safeguard personal information that is without question among the most sensitive and most deserving of protection. Considering also that the state compels the collection of this information for public safety and research purposes, the subsequent disclosure for other unrelated purposes implicates a Constitutional interest in informational privacy.

## ARGUMENT

### **I. Medical Privacy is a Fundamental Concern for Patients.**

There are approximately 1.4 million health care providers in the United States. *IMS Health Inc. v. Ayotte*, 490 F. Supp. 2d 163, 165 (D.N.H. 2007), *rev'd and vacated*, 550 F.3d 42 (1st Cir. 2008). These providers write billions of prescriptions each year for more than 8,000 different pharmaceutical products. *Id.* These prescriptions are filled at 54,000 retail pharmacies throughout the country. *Id.* The retail pharmacies acquire records for every prescription they fill. These records include: patient name; prescriber identification; drug name; dosage prescribed; quantity; and date the prescription was filled. *Id.* In order to comply with federal and state privacy laws, patient-identifying information is obscured through a cryptographic technique and the record deidentified, often with software installed by the data-mining companies themselves. *Id.* at 166. The rest of the prescription record remains intact. Thus, a patient's entire drug history is correlated, and each provider can be identified along with his or her prescribing habits. Each provider is individually identified in the data, along with his or her entire prescribing history broken down by patient. *This* is the very valuable information that pharmacies are selling to data-mining companies. *Id.* This practice raises privacy concerns for both patients and health care providers.

Public sentiment overwhelmingly favors the protection of patient privacy. Over 70% of Americans have concerns over the disclosure of their medical information without their knowledge. Harris

Interactive, *HIPAA Notices Have Improved Public's Confidence That Their Medical Information is Being Handled Properly: However public split on benefits of and privacy risks associated with Electronic Medical Records (EMR)*, Feb. 24, 2005.

Arizona, the District of Columbia, Illinois, Kansas, Maine, Maryland, Massachusetts, Nevada, New Hampshire, New York, North Carolina, Rhode Island, Texas, Washington, and West Virginia have all considered or enacted bills banning the sale of prescriber data. Ariz. Rev. Stat. Ann. § 32-1973 (2010); D.C. Code § 3-1207.41 (2010); H.B. 1459, 95th Gen. Assem. (Ill. 2007); S.B. 229 (Kan. 2007); Me. Rev. Stat. Ann. tit. 22, § 1711-E (2010); S.B. 266 (Md. 2007); S.B. 1275 (Mass. 2007); S.B. 231 (Nev. 2007); N.H. Rev. Stat. Ann. §§ 318:47-f, 318-47g, 318-B:12 (2010); H.B. 5891B, Reg. Sess. (N.Y. 2009); S.B. 159, Gen. Assem. (N.C. 2007); S.B. 2683, Gen. Assem. (R.I. 2008); S.B. 1620 (Tex. 2007); H.B. 1850 (Wash. 2008); W. Va. Code. § 30-5-12c (2010); *see also* Joe Mullin, *States Consider Limits on Medical Data-mining*, BOSTON GLOBE, Apr. 7, 2007; The Prescription Project, *Prescription Data Mining Fact Sheet*, Nov. 19, 2008.<sup>3</sup> Vermont's prescription privacy law, the focus of this case, does not ban the sale of prescriber data. It merely gives health care providers the choice to keep their data private. *See* Vt. Stat. Ann. tit. 18, § 4631 (2010).

Doctors have also petitioned the American Medical Association ("AMA") on behalf of themselves and their patients for legal relief, blaming data-

---

<sup>3</sup>[http://www.prescriptionproject.org/tools/fact\\_sheets/files/0004.pdf](http://www.prescriptionproject.org/tools/fact_sheets/files/0004.pdf).

mining companies for interfering with the patient-doctor relationship and violating doctors' and patients' privacy. Tanya Alberts, *Doctors Ask AMA to Assure Some Privacy for their Prescription Pads*, AMNews, Dec. 25, 2000. Even after the AMA adopted an opt-out approach to the sale of prescriber data, doctors continued to question the practice of selling prescriber data and lobbied for stronger safeguards for patient information. Joe Mullin, *States Consider Limits on Medical Data-mining*, Boston Globe, Apr. 7, 2007.

Although the AMA's Prescribing Data Restriction Program ("PDRP") allows physicians to opt out of having their prescribing history accessed by drug representatives, many physicians believe it is inadequate. The National Physician's Alliance supports a complete ban on the sale of prescriber data. Nat'l Physician's Alliance, *Issue Brief: The Sale of Physician Prescribing Data Raises Health Care Costs* (Feb. 2009).<sup>4</sup> They have spoken against the PDRP because the program is burdensome and not widely publicized.

Health care providers face the unique challenge of providing quality, affordable health care, while protecting each patient's fundamental right to privacy. The use of electronic databases reduces institutional costs, integrating applicable data from multiple sources, and allowing patients to receive a higher and more accurate level of care, but without proper safeguards, these databases pose a serious

---

<sup>4</sup> Available at [http://npalliance.org/images/uploads/IssueBrief-Prescribing\\_Data\\_low\\_res.pdf](http://npalliance.org/images/uploads/IssueBrief-Prescribing_Data_low_res.pdf).

threat to privacy. *See* Latanya Sweeney, *Weaving Technology and Policy Together to Maintain Confidentiality*, 25 J. LAW, MED., & ETHICS 98, 98-99 (1997). (summarizing industry and research use of personally identifiable health care information). However, this transition to centralized depositories for health care information may lead to the disclosure of private medical records to secondary actors, such as researchers, economists, statisticians, administrators, consultants, and computer scientists.

## **II. The Vermont Law Seeks to Safeguard Medical Privacy, a Fundamental Concern for those Receiving Prescription Drugs**

The state of Vermont enacted the Confidentiality of Prescription Information law, Vt. Stat. Ann. tit. 18, § 4631 (2010), to address concerns about the privacy of medical information. The Vermont legislature found a “reasonable expectation that the information in the prescription, including [the doctor’s] identity and that of the patient, will not be used for purposes other than the filling and processing of the payment for . . . prescription[s].” Vt. Acts & Resolves 80 (S.115) § 1(29).<sup>5</sup> This legislative determination followed from the Vermont Medical Society’s 2006 Medical Privacy Resolution opposing data-mining practices. C.A. App. A4197, Vermont Medical Society Resolution, “Ensuring the Privacy of Prescription Information,” Oct. 14, 2006, (“the doctor-patient relationship requires confidentiality and privacy to work effectively.”); *see also* Stolberg & Gerth, *High-Tech*

---

<sup>5</sup> <http://www.leg.state.vt.us/docs/legdoc.cfm?URL=/docs/2008/acts/ACT080.htm>

*Stealth Being Used to Sway Doctor Prescriptions*, N.Y. TIMES, Nov. 16, 2000 at A1 (Georgetown University Law Center Professor Lawrence O. Gostin describes patient profiling as a fundamental violation of privacy) *cited* at C.A. App. A4218-22.

Vermont's concern is understandable. Doctors in Vermont and across the country are required under state, federal, and international law to retain their patients' personal information and transmit it to the pharmacies that service their prescriptions. *See, e.g.*, Vt. Bd. Of Pharmacy Admin. Rule § 9.1, 9.24, 9.26, 9.27; 21 U.S.C. § 822, 827 (2010); Single Convention on Narcotic Drugs, art. 19, 20, Mar. 30, 1961, 18 U.S.T. 1407, 520 U.N.T.S. 151; Convention on Psychotropic Substances, art. 16, opened for signature Feb. 21, 1971, 1019 U.N.T.S. 175 (ratified by the United States in 1980). Data retention laws require that pharmacies obtain from doctors the full name, street address, age, and gender of their patients, as well as the name, strength, dosage, and number of refills of their prescribed drugs. *See, e.g.*, Vt. Bd. Of Pharmacy Admin. Rule § 9.1, 9.24.

Federal law requires pharmacies to remove names and addresses from this information, in addition to month and day, but not year, of birth, and encrypt the rest before selling it. Privacy Rule of the Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub. L. No. 104-191 (1996), 45 C.F.R. §§ 164.312(e)(2)(ii), 164.514(b)(2)(i) (2010). Federal law imposes no additional restrictions on pharmacies and companies such as IMS Health and Verispan that sell this information. 45 C.F.R. § 164.502(d)(2) (2010). IMS Health will sell the information "to anyone who wants to buy it." C.A. App. A78 (trial testimony of Hossam Sadek, General

Manager for IMS Health's Business Line Management).

It is the provision of the Act that permits marketing use of "patient and prescriber data" that "does not identify a prescriber and [for which] there is no reasonable basis to believe that the data provided could be used to identify a prescriber" that gives rise to EPIC's brief. 18 Vt. Stat. Ann. tit. 18, § 4631 (2010). Simply stated, the privacy interest that undergirds the state's interest in this statute is even greater than what the legislature expressly recognized in the findings.

IMS Health's deidentification techniques do not adequately protect the privacy interests of patients or preserve doctor-patient confidentiality. For this reason, the Court should give even greater weight to patients' privacy interests in its *Central Hudson* analysis, if the Court determines that the transfer of nonpublic patient data is in fact commercial speech. *See Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n of N.Y.*, 447 U.S. 557 (1980); *see also IMS Health v. Ayotte*, 550 F.3d 42, 51-53 (1st Cir. 2008), *cert. denied*, 129 S. Ct. 2864 (2009) (the commercial use of nonpublic information is better described as commercial conduct than commercial speech.)

The patient interest in protecting the privacy of personal medical information is widely recognized. "A majority of adults express discomfort (42 percent) or uncertainty (25 percent) with their health information being shared with other organizations—even if . . . [their] name, address, [date of birth, and social security number] were not included." California Healthcare Foundation, *Consumers and*



*Health Information Technology: A National Survey*, 26 (2010).<sup>6</sup> One out of every seven adults “would hide something from their doctor if they knew their information would be shared,” even with guarantees that their names, addresses, dates of birth, and social security numbers stay secret. *Id.* at 25. Another one-third “would consider hiding information.” *Id.* Over 90% of Americans want to determine which companies and government entities can see their health information. *Patient Privacy Rights/Zogby International Poll*, Nov. 23, 2010.

Patients have a privacy interest in confidentiality of treatment. *See Discussion Draft of Health Information Technology and Privacy Legislation: Hearing on H.R. 6357 Before the Subcomm. on Health of the H. Comm. on Energy and Commerce*, 110<sup>th</sup> Cong. (2008) (Testimony of Dr. Deborah Peel); Anita Allen, *Privacy and Medicine*, Stanford Encyclopedia of Philosophy (2009). The confidentiality interest complements the patients’ interest in ensuring that personal data remains anonymous. Confidentiality of treatment encourages patients to seek the most accurate, and therefore best possible, care by promoting a trusting and frank relationship between patient and doctor. *See id.* Statutes that safeguard doctor-patient confidentiality also advance a First Amendment interest in protecting the ability of individuals to freely express their views to one another. “These statutes undeniably protect this venerable right of privacy.

---

<sup>6</sup> available at <http://www.chcf.org/~media/Files/PDF/C/PDF%20ConsumersHealthInfoTechnologyNationalSurvey.pdf>.

Concomitantly, they further the First Amendment rights of the parties to the conversation.” *Bartnicki v. Vopper*, 532 U.S. 514, 553 (2001) (Rehnquist, C.J., dissenting). There is an additional First Amendment interest in protecting the right of the individual to not speak publicly if he or she chooses not to do so. *Wooley v. Maynard*, 430 U.S. 705 (1977). Privacy laws that allow individuals to selectively disclose communications on private matters to others advance this important Constitutional interest. *See, e.g., Nat’l Cable & Telecomms. Ass’n v. FCC*, 555 F.3d 996 (D.C. Cir. 2009).

Recognition of this confidentiality interest dates back to the Hippocratic Oath: “I will respect the privacy of my patients, for their problems are not disclosed to me that the world may know.” This portion of the Oath was originally designed to reflect the “Hippocratic bargain,” wherein patients “relinquished aspects of their privacy in exchange for their physicians’ assurances of confidentiality.” Mark Rothstein, *The Hippocratic Bargain and Health Information Technology*, 38 J. LAW, MED., & ETHICS 11 (2010).

Confidentiality also protects the doctor-patient relationship from external influence. “Medical confidentiality promotes medical autonomy by sheltering those seeking controversial medical care from outside criticism and interference with decisions.” Anita Allen, *Privacy and Medicine*, Stanford Encyclopedia of Philosophy (2009).<sup>7</sup>

Even if the technique to separate the data from an identifiable individual is adequate, the transfer of

---

<sup>7</sup> <http://plato.stanford.edu/entries/privacy-medicine/>.

sensitive medical information may still implicate a cognizable privacy interest. Individuals have an “interest in the uses to which data sets that include their data is put, even if they are not personally identified by researchers.” *Id.*<sup>8</sup> Professor Jerry Kang makes clear the privacy interest that remains even after identifying information has been eliminated:

Imagine that a psychiatrist publishes verbatim counseling notes in a best-selling book, but in a way that the specific identity of the patient cannot be determined. If the patient protests at having her story chronicled in agonizing detail to the public, could the good doctor respond that because the information is not identifiable to the specific patient, even with additional research, it is not “personal information.” And, because it is not personal information, the patient lacks any privacy claim? To my mind, this reasoning fails to account for the residual privacy interest that exists, notwithstanding the anonymity.

Jerry Kang, *Cyberspace Privacy*, 50 STAN. L. REV. 1193, 1209 (1998). There are “non-medical harms associated with the excessive disclosure of health information, including embarrassment, strains on intimate relationships, stigmatization, and discrimination.” Mark Rothstein, *The Hippocratic Bargain and Health Information Technology*, 38 J. LAW, MED., & ETHICS 11 (2010). As Judge Posner

---

<sup>8</sup> <http://plato.stanford.edu/entries/privacy-medicine/>.

wrote for the Seventh Circuit in *Northwestern Memorial Hospital v. Ashcroft*, 362 F.3d 923 (7th Cir. 2004) 45 C.F.R. § 164.502(d)(2).involving access to redacted medical records:

Even if there were no possibility that a patient's identity might be learned from a redacted medical record, there would be an invasion of privacy. Imagine if nude pictures of a woman, uploaded to the Internet without her consent though without identifying her by name, were downloaded in a foreign country by people who will never meet her. She would still feel that her privacy had been invaded. The revelation of the intimate details contained in the record of a late-term abortion may inflict a similar wound.

The transfer of sensitive prescription information implicates a range of privacy interests.

### **III. Data-Mining Companies Fail to Mitigate Medical Privacy Risks to Vermont Residents**

The protection of the patient privacy interest in the transfer of prescriber information from the pharmacy to the data-mining firm relies upon two distinct techniques, both of which are inadequate to the task at hand.

The first is the deidentification of the patient's actual identity through a cryptographic technique known as "hashing." In ideal circumstances, a record containing the hashed representation of the patient's actual identity could never be linked to the actual patient. But the cryptographic technique chosen to

protect patient privacy in this matter has been suspect for at least 15 years, can now be broken using nothing more than an ordinary desktop computer, and is considered unsuitable for further use by the federal government. Vlastimil Klima, *Finding MD5 Collisions – A Toy For a Notebook* (Mar. 5, 2005);<sup>9</sup> Chad Dougherty, *Vulnerability Note VU#836068: MD5 Vulnerable to Collision Attacks*, United States Computer Emergency Readiness Team (Dec. 31, 2008).<sup>10</sup> See generally, Wikipedia, “MD5,” (“ . . . it has been shown that MD5 is not collision resistant; as such, MD5 is not suitable for applications like SSL certificates or digital signatures that rely on this property. . . . The security of the MD5 hash function is severely compromised.”)<sup>11</sup>

The second technique is the reduction in data elements in the record to reduce the likelihood that the actual identity could be inferred from other information such as age, gender, and home address. The records disseminated by IMS Health contain, in addition to detailed drug and prescriber information, the age and gender, but not the home address, of the patient. This, in combination with other publicly available databases, is enough information to reidentify patients.

IMS Health asserts that “there is no way that you can actually reverse engineer the data back to a patient.” C.A. App. A89 (trial testimony of Hossam

---

<sup>9</sup> [http://cryptography.hyperlink.cz/md5/MD5\\_collisions.pdf](http://cryptography.hyperlink.cz/md5/MD5_collisions.pdf).

<sup>10</sup> <http://www.kb.cert.org/vuls/id/836068>.

<sup>11</sup> <http://en.wikipedia.org/wiki/MD5> (accessed on Feb. 28, 2011).

Sadek, General Manager for IMS Health's Business Line Management). IMS's confidence is misplaced. The hash technique the company uses is no longer considered reliable by the scientific community or the federal government, and the data elements that remain in the record make reidentification, particularly in a small state such as Vermont, relatively simple. These factors buttress the state's interest in limiting the transfer of prescriber information. "A company falsely believing the data could not be re-identified may unknowingly put data at risk or not seek necessary security precautions. After all, data that adheres to the HIPAA Safe Harbor or Scientific Standard provisions can be shared freely without HIPAA review and sanctions." 45 C.F.R. § 164.502(d)(2).

As described below, there are significant shortcomings in the cryptographic and deidentification techniques adopted by Respondent to protect prescriber information. These shortcomings expose the patient information IMS Health and Verispan collect to the risk of actual disclosure and further underscore the privacy interest at issue in this case. Striking down Vermont's confidentiality statute, which allows physicians to limit the disclosure of prescriber information, would make this inadequately protected medical information more widely available.

***A. The Cryptographic Technique Used to Conceal the Identity of Patients is Inadequate***

Verispan uses the MD5 Hash Algorithm to conceal the actual identity of patients who receive prescription medications. C.A. App. A99 (trial

testimony of Jody Fisher, Vice President of Verispan's Product Management); *see also* 45 C.F.R. § 164.312(e)(2)(ii) (2010). MD5 was developed by Ron Rivest in 1991. MD5 is a cryptographic "hash function" that creates a fixed length "digest" based on a text input. As such, it is possible to transform a person's name into a unique code and, in theory, not to determine the original name from the resulting code. Ron Rivest, *The MD5 Message-Digest Algorithm*, RFC 1321 (Apr. 1992).<sup>12</sup> MD5 is an improved version of MD4 and is similar in design. BRUCE SCHNEIER, *APPLIED CRYPTOGRAPHY* 436 (2nd ed. 1996).

A plainly inadequate transformation would be based on ROT13 ("rotate by 13 places"), a simple substitution cipher that traces its roots to Julius Caesar. *See generally id.* at 11; Wikipedia, "ROT13." With ROT13 each letter in a string is replaced by the letter 13 characters further along in the alphabet. For example, a ROT13 transformation of "Alan Turing" would produce "Nyna Ghevat." If someone were to examine a patient record with the string "Nyna Ghevat" appearing where a "deidentified" name would be expected, it would not be difficult to determine the actual name that generated the string. Hashing algorithms are typically far more sophisticated, and will for example routinely produce a fixed-length output regardless of the length of the initial text. But this example with ROT13 makes clear the risk that an inadequate technique for deidentification establishes an ongoing privacy risk.

---

<sup>12</sup> <http://www.faqs.org/rfcs/rfc1321.html>.

For 15 years, security expert Bruce Schneier has been “wary of using MD5” because of analytic work proving MD4 and MD5 had security vulnerabilities. *Id.* at 441. Two different teams of analysts demonstrated that MD5’s underlying cryptographic key algorithm was insufficiently random. *Id.* Randomness is an important metric. “The algorithms that take a block of data and hide it in the noise . . . need data that is as close to random as possible. This lowers the chance that it can be detected.” Peter Wayner, *DISAPPEARING CRYPTOGRAPHY* 32 (2nd ed. 2002). The second team partially compromised MD4, and then demonstrated that MD5 failed by its own design principles. Bert den Boer and Antoon Bosselaers, *Collisions for the Compression Function of MD5*, Proceedings of Eurocrypt '92, Advances in Cryptology, 71-88 (1992).

Researchers in China in 2004 and the Czech Republic in 2005 moved beyond analytical work into demonstrated applications of MD5’s vulnerabilities. Both successfully compromised MD5, using “just ordinary desktop computers” rather than supercomputers. Vlastimil Klima, *Finding MD5 Collisions – A Toy For a Notebook* (Mar. 5, 2005).<sup>13</sup>

In December of 2005, Ron Rivest declared MD5 “clearly broken.” Ron Rivest, *[Python-Dev] hashlib - faster md5/sha, adds sha256/512 support* (Dec. 16, 2005).<sup>14</sup> Significantly, the government agency charged with safeguarding federal computer systems

---

<sup>13</sup>

[http://cryptography.hyperlink.cz/md5/MD5\\_collisions.pdf](http://cryptography.hyperlink.cz/md5/MD5_collisions.pdf)

<sup>14</sup> <http://mail.python.org/pipermail/python-dev/2005-December/058850.html>



determined that the technique upon which Respondent relies to safeguard the actual identity of patients is no longer reliable. The Department of Homeland Security's Computer Emergency Readiness Team concluded that MD5 is "cryptographically broken and unsuitable for further use." Chad Dougherty, *Vulnerability Note VU#836068: MD5 Vulnerable to Collision Attacks*, United States Computer Emergency Readiness Team (Dec. 31, 2008).<sup>15</sup> Bruce Schneier added that "no one should be using MD5 anymore." Bruce Schneier, *Forging SSL Certificates*, Schneier on Security (Dec. 31, 2008).<sup>16</sup>

The legal consequence of the problem with MD5 is to strengthen the state's interest in limiting the transfer of prescriber data. The deidentification technique deployed by Respondent's will continue to leave at risk the disclosure of actual patient prescription data.

### ***B. Patient Records are At Risk of Being Reidentified***

*Amici* EPIC believe that "deidentified" data "should not be considered anonymous." See Latanya Sweeney, *Simple Demographics Often Identify People Uniquely* (Carnegie Mellon University, Data Privacy Working Paper No. 3, 2000);<sup>17</sup> see also Ross Anderson, *The DeCODE Proposal for an Icelandic*

---

<sup>15</sup> <http://www.kb.cert.org/vuls/id/836068>.

<sup>16</sup> [http://www.schneier.com/blog/archives/2008/12/forging\\_ssl\\_cer.html](http://www.schneier.com/blog/archives/2008/12/forging_ssl_cer.html)

<sup>17</sup> <http://dataprivacylab.org/projects/identifiability/paper1.pdf>

*Health Database*, at 11, Oct. 20, 1998 (“The use of a code to replace identifiers is in any case not sufficient to secure anonymity.”)<sup>18</sup> There are bits of patient data in “deidentified” medical records, called “quasi-identifiers” that link the medical information in any given deidentified record back to a “small and limited set of” real world subjects. *Id.*

Record linkage is a technique that reconstitutes “deidentified” records by matching quasi-identifiers with databases of other publicly accessible records (*i.e.*, “identification databases”). Fida Kamal Dankar & Khaled El Emam, *A Method for Evaluating Marketer Re-identification risk*, *PROCEEDINGS OF THE 2010 EDBT/ICDT WORKSHOPS*, ACM, Article 28 (2010);<sup>19</sup> Khaled El Emam *et al.*, *Evaluating Common De-identification Heuristics for Personal Health Information*, 8 *J. MED. INTERNET RES.* 4 (2006). Record linkages can be established between hospital discharge records and demographic data from the 1990 US Census. Latanya Sweeney, *Simple Demographics Often Identify People Uniquely* (Carnegie Mellon University, Data Privacy Working Paper No. 3, 2000).<sup>20</sup> “In principle an identification database can be constructed in a number of ways,” including aggregating data from public registries such as voter lists, contacting commercial organizations that sell data about members of the general public, and collecting information online from

---

<sup>18</sup> Available at

<http://www.cl.cam.ac.uk/~rja14/Papers/iceland.pdf>.

<sup>19</sup> <http://portal.acm.org/citation.cfm?id=1754271>

<sup>20</sup> <http://dataprivacylab.org/projects/identifiability/paper1.pdf>

individuals who maintain profiles, post their *curriculae vitae*, or publish personal web pages. Khaled El Emam *et al.*, *Evaluating Common De-identification Heuristics for Personal Health Information*, 8 J. MED. INTERNET RES. 4 (2006).

Almost all medical patients can be reidentified using the zip code, date of birth, and gender categories on their deidentified records. Latanya Sweeney, *K-anonymity: A Model for Protecting Privacy*, 10 INTERNATIONAL JOURNAL ON UNCERTAINTY, FUZZINESS AND KNOWLEDGE-BASED SYSTEMS, no. 5, 557-70 (2000).<sup>21</sup> One source for this information is IMS Health and Verispan themselves; they buy and sell medical records which contain date of birth, gender, and zip code information for millions of patients. See 45 C.F.R. § 164.514(b)(2)(i)(B)-(C) (2010). With the prescription records they buy from pharmacies, they create unique identifiers for each patient so that deidentified records link to one another.

What we do is encrypt the information, strip out all of the identifiable information, and replace it with the serial linking code. That linking code is several digits long. It's about in its native form about 39 digits long actually and what we do is we strip off the information, replace it with this linking code, so that every time an entity comes into the data base, it's replaced with the same code. So you can follow an

---

<sup>21</sup> <http://dataprivacylab.org/dataprivacy/projects/kanonymity/kanonymity.pdf>

individual over time, but you have no idea who that individual actually is.

C.A. App. A99 (trial testimony of Jody Fisher, Vice President of Verispan's Product Management).

Federal law does *not* require these companies to remove from their records the year of the patient's birth, the patient's gender, or the identities and official addresses of the patient's pharmacist and doctor. 45 C.F.R. § 164.502(d)(2); *IMS Health Inc. v. Sorrell*, 630 F.3d 263 (2d Cir. 2010) n.4 (Livingston, J., dissenting), *cert. granted*, 131 S. Ct. 857 (U.S. 2011) (No. 10-779) (describing a deidentified record: "50-year-old woman who lives in Central Vermont; has prescriptions filled in Montpelier; [and] is a patient of Dr. Jones in Montpelier ... regularly takes an antidepressant and a cholesterol-lowering drug"). There are better technical and legal solutions to safeguard identity. "Differential privacy" is a more robust standard for ensuring the medical data is not exposed. Latanya Sweeney, *Patient Identifiability in Pharmaceutical Marketing Data*, Cambridge, Data Privacy Working Paper No. 1015 (2011) at 19.<sup>22</sup> Instead of seeking the goal of deidentifying data, differential privacy "formally defines" what it means for data practices to be "privacy-preserving." Arvind Narayanan and Vitaly Shmatikov, *Myths and Fallacies of "Personally Identifiable Information,"* 53 COMM'N OF THE ACM, 24-26 (June 2010); *see also* David Chaum, *Achieving Electronic Privacy*, SCI. AM., 96-101 (1992).

---

<sup>22</sup><http://dataprivacylab.org/projects/identifiability/pharma1.pdf>

There are additional concerns arising from the collection of patient data over time as is clearly contemplated by Respondents. As Professor Ross Anderson has explained:

Firstly, although it is not too difficult to de-identify data that provide only a time-limited snapshot of a population's health – such as the data which health services use to compile monthly management statistics of numbers of operations, consumption of drugs and the like – it is effectively impossible to de-identify longitudinal records, that is, records which link together all (or even many) of the health care encounters in a patient's life.

Ross Anderson, *The DeCODE Proposal for an Icelandic Health Database*, at 3, Oct. 20, 1998 (“The use of a code to replace identifiers is in any case not sufficient to secure anonymity.”)

The Institute of Medicine (“IOM”) has recently sought to address the problem that “deidentified” data is invariably an imperfect technique and legal steps should be taken to safeguard the underlying privacy interest. The IOM has concluded that “unauthorized re-identification of information that has had direct identifiers removed should be prohibited by law, and violators should face legal sanctions.” BEYOND THE HIPAA PRIVACY RULE: ENHANCING PRIVACY, IMPROVING HEALTH THROUGH RESEARCH, 265 (Sharyl J. Nass et al, ed., 2009).<sup>23</sup> The IOM has stated that “unauthorized re-identification of information that has had direct identifiers removed should be prohibited by law, and

---

<sup>23</sup> [http://www.nap.edu/catalog.php?record\\_id=12458](http://www.nap.edu/catalog.php?record_id=12458)

violators should face legal sanctions.” *Id.* at 265. In addition, the IOM determined, “researchers receiving information with direct identifiers removed should be required to establish security safeguards and to set limits on access to data.” *Id.*

### ***C. Other Data Is At Risk of Being Reidentified***

In addition to the problems associated with the use of inadequate techniques to deidentify subjects in a prescriber record database, there is also the privacy risk that the identity may be inferred from related data elements. As Professor Ross Anderson explains the “basic problems of inference Control in Medicine:”

The standard way of protecting such information is to remove patients’ names and addresses from their records, and thus make them anonymous. But this is rarely sufficient. If a database allows detailed enough queries, then individuals can still be identified, and this is especially so if information about different clinical episodes can be linked.

ROSS J. ANDERSON, SECURITY ENGINEERING: A GUIDE TO BUILDING DEPENDABLE DISTRIBUTED SYSTEMS 172 (2008).

The National Institutes of Health has come to acknowledge that deidentification is an insufficient method of ensuring patient privacy, and therefore limits the accessibility of its deidentified genomic data to principal medical researchers. Homer N, *et al.*, *Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using*

*High-Density SNP Genotyping Microarrays*. 4 PLOS GENETICS, no.8 (August 2008);<sup>24</sup> Rachel Ehrenberg, *Hiding Patients in Plain Sight*, Science News, Apr. 12, 2010; Matthew D. Mailman, et al., *The NCBI dbGaP Database of Genotypes and Phenotypes*, 39 NATURE GENETICS, 1181 (2007).<sup>25</sup> “Even when de-identified, [this data will] remain unique to the individual and could potentially be linked to a specific person if used in conjunction with other databases.” *Id.*

Genetic researchers at NIH correlate large pools of genetic data with large pools of clinical data in order to find and track links between specific genes and physical traits and diseases. See Grigorios Loukides, et. al., *Anonymization of Electronic Medical Records for Validating Genome-wide Association Studies*, 107 PROCEEDINGS OF THE NAT'L ACAD. OF SCI., no. 17, 7898, (March 11, 2010).<sup>26</sup> NIH conserves this data instead of discarding it, storing it in “repositories for re-use.” Matthew D. Mailman, et al., *The NCBI dbGaP Database of Genotypes and Phenotypes*, 39 NATURE GENETICS, 1181-86 (2007). The data is deidentified. *Id.*

NIH’s previous practice was to make the deidentified information available “to anyone with Internet access.” Rachel Ehrenberg, *Hiding Patients*

---

<sup>24</sup><http://www.plosgenetics.org/article/fetchObjectAttachment.action?uri=info%3Adoi%2F10.1371%2Fjournal.pgen.1000167&representation=PDF>

<sup>25</sup><http://www.nature.com/ng/journal/v39/n10/full/ng1007-1181.html>

<sup>26</sup><http://www.pnas.org/content/early/2010/04/05/0911686107.full.pdf+html>

in *Plain Sight*, Science News, Apr. 12, 2010. In 2008, researchers officially demonstrated that repositories of deidentified genetic data are vulnerable to the same reidentification techniques Dr. Latanya Sweeney identified in prescriber records. Homer N, *et al.*, *Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays*. 4 PLOS GENETICS, no.8 (Aug. 2008). This demonstration prompted NIH to restrict access to its deidentified information. Rachel Ehrenberg, *Hiding Patients in Plain Sight*, Science News, Apr. 12, 2010.

Furthermore, the risk of reidentification is influenced by collections of information extending beyond medical data.

Data sufficiently de-identified today may be re-identifiable tomorrow because there is no knowledge or coordination of datasets that may be available tomorrow. . . [A]s more data is made readily available, such as credit card purchases, online prescription purchases, email messages about refills, and cell phone location data, re-identification risks increase.

Latanya Sweeney, *Patient Identifiability in Pharmaceutical Marketing Data*, Cambridge, Data Privacy Working Paper No. 1015 (2011) at 18.

Researchers have been able to reidentify anonymized records in databases as varied as Social Security records, Internet search queries, and video rentals. For example, researchers at Carnegie Mellon University developed a reidentification process that makes “statistical inference[s]” about SSNs based on a person’s birth date and publically available



information from the Social Security Administration about how it assigns SSNs. Alessandro Acquisti and Ralph Gross, *Predicting Social Security Numbers from Public Data*, 106 PROCEEDINGS OF THE NAT'L ACAD. OF SCI., no. 27, 10975 (Jul. 7, 2009).<sup>27</sup> The researchers suggested that attackers can “exploit online services [such as instant credit approval services, mass phishing emails, or the SSA’s own SSN Verification Service] as oracle machines” in order to verify correlations between SSN and birth date.” *Id.* Their ability to predict SSNs increased with the presence of unique identifiers for younger individuals and those who lived in less-populous states. *Id.*

Another example is AOL’s release of anonymous public search records of 20 million Internet search queries from 657,000 people made over a three-month period. Michael Barbaro & Tom Zeller Jr., *A Face is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES, Aug. 9, 2006, at A1. The New York Times published its own work reidentifying supposedly anonymous records, linking search queries back to the individual who made them. *Id.* Bloggers set up websites to make it easier for the public to search AOL’s data and reidentified additional search records. *Id.* “*Paul Ohm, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization 4 (University of Colorado Law Legal Studies Research Paper No. 09-12, 2009*

Finally, researchers from the University of Texas found that if an adversary knows six precise ratings a

---

<sup>27</sup> <http://www.pnas.org/content/early/2009/07/02/0904891106.full.pdf+html>

person in the Netflix video-rental database has assigned to obscure movies, without any other information, the adversary can identify that person 84% of the time. Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Datasets (How to Break the Anonymization of the Netflix Prize Dataset)*, PROC. OF 29TH IEEE SYMPOSIUM ON SEC. AND PRIVACY, Oakland, CA, 111-125 (May 2008). When the researchers included the specific times when ratings assigned, they successfully identified 99% of the people in the Netflix database. *Id.*

***D. Overturning the Vermont Statute Poses a Risk of Increasingly Widespread Access to Health Data***

Not only does the Vermont confidentiality law seek to safeguard an important interest in the protection of medical information, a decision to overturn the statute could make it exceedingly difficult to address harms when they occur. Dr. Sweeney has warned that “[a] person could be egregiously harmed by data sharing, but not be able to show the hidden trail that led to the harm. An example is the compilation and use of personal prescription profiles by companies.” Latanya Sweeney, *Patient Identifiability in Pharmaceutical Marketing Data*, Cambridge, Data Privacy Working Paper No. 1015 (2011) at 19.

Proprietary tracking technology captures medically related search engine queries. *See e.g.*, Marketing Technology Solutions Corp. (MTS), *QH*

*Connect*.<sup>28</sup> Data-mining companies match those queries with outside databases and compile the resulting information into personal medical profiles for sale to pharmaceutical companies. *See, e.g., AOL Advertising, Case Study: OTC Pharma Leader Drives Offline Sales with AOL's Online Targeting*,<sup>29</sup> Data-mining companies have populated millions of individual profiles with billions of tracked online health information queries. Demand Media, Inc., “Form S-1,” August 6, 2010, p104; MTS, *QH Connect*.<sup>30</sup> Their profiles detail each patient’s health conditions, preferred treatments, doctor relationships, plans to visit the doctor, household-level purchasing history, locational data, insurance claims data, and up to 250 other personal data points. MTS *QH Connect*,<sup>31</sup> Quality Health, *Privacy Policy*.<sup>32</sup>

If states cannot regulate these practices, the only applicable protections will be the same federal deidentification techniques which researchers have revealed as inadequate across a number of industries. *See* 45 C.F.R. §§ 164.312(e)(2)(ii), 164.514(b)(2)(i) (2010); Razorfish LLC, *Outlook Report 2010* at 35.<sup>33</sup> (“This data includes HIPAA-compliant medical claim data that is stripped of personally-identifiable

---

<sup>28</sup> <http://www.qualityhealth.com/privacyPolicy/footer>

<sup>29</sup> <http://advertising.aol.com/sites/default/files/OTC-targeting.pdf>

<sup>30</sup> [http://stage.mtscorp.com/qh\\_connect.html](http://stage.mtscorp.com/qh_connect.html)

<sup>31</sup> [http://stage.mtscorp.com/qh\\_connect.html](http://stage.mtscorp.com/qh_connect.html)

<sup>32</sup> <http://www.qualityhealth.com/privacyPolicy/footer>

<sup>33</sup> <http://razorfishoutlook.razorfish.com/publication/?m=11995&1=1>.

information, and targets selected condition sufferers down to the ZIP code+4 geographic level.”).

Computer scientists understand that techniques can be developed to safeguard privacy and protect identity, but this does not obviate the need for legal protections that recognize the importance of privacy. As Jerome Weisner, former President of MIT and the first Science Advisor to the President, explained:

There are those who hope new technology can redress these invasions of personal autonomy that information technology now makes possible, but I don't share this hope. To be sure, it is possible and desirable to provide technical safeguards against unauthorized access. It is even conceivable that computers could be programmed to have their memories fade with time and to eliminate specific identity. Such safeguards are highly desirable, but the basic safeguards cannot be provided by new inventions. They must be provided by the legislative and legal systems of this country. We must face the need to provide adequate guarantees for individual privacy.<sup>34</sup>

The Vermont Prescription Confidentiality Law seeks to address this interest.

---

<sup>34</sup> *Federal Data Banks, Computers and the Bill of Rights: Hearings Before the Subcomm. on Constitutional Rights of the House Comm. on the Judiciary*, 92d Cong., 1st Sess. Part I, 761-74 (1971) (testimony of Jerome B. Wiesner, provost elect, Massachusetts Institute of Technology).

#### IV. Data-mining of Prescriber Information Implicates the Constitutional Right to Informational Privacy

The mandatory collection of personal medical information by the state, coupled with the risk of subsequent disclosure, implicates privacy interests of a Constitutional dimension. In *NASA v. Nelson*, 131 S. Ct. 746 (2011), the Court recently acknowledged that *Whalen v. Roe*, 429 U.S. 589 (1977) upheld a right to informational privacy rooted in the Constitution. 131 S.Ct. at 763. Here, a government-mandated medical data retention regime will expose individuals to a violation of that right. The specific concern about the adequacy of safeguards to protect the privacy of medical record information should bear directly on the Court's analysis.

In *Whalen*, the Court considered “whether the State of New York may record, in a centralized computer file, the names and addresses of all persons who have obtained, pursuant to a doctor's prescription, certain drugs . . .” 429 U.S. at 591. The Court catalogued the tiers of robust legislative and physical data protections ensuring that individual technological failures would not result in the exposure of personal information. The Court noted that “The receiving room is surrounded by a locked wire fence and protected by an alarm system.” *Id.* at 594. It further observed that “[t]he computer tapes containing the prescription data are kept in a locked cabinet” and that “[w]hen tapes are used, the computer is run ‘off-line,’ which means that no terminal outside of the computer room can read or record any information.” *Id.* The Court also pointed out that “[p]ublic disclosure of the identity of patients

is expressly prohibited by the statute and by a Department of Health regulation. Willful violation of these prohibitions is a crime punishable by up to one year in prison and a \$2,000 fine.” *Id.* at 594-95.

In concurrence, Justice Brennan further noted, “[t]he central storage and easy accessibility of computerized data vastly increase the potential for abuse of that information, and I am not prepared to say that future developments will not demonstrate the necessity of some curb on such technology.” *Id.* at 607 (Brennan, J., concurring). Justice Brennan ultimately agreed with the Court’s majority, observing that “the State’s carefully designed program includes numerous safeguards intended to forestall the danger of indiscriminate disclosure.” *Id.* at 607.

In *NASA*, the Court cited this rationale as a principle of law, holding “the mere possibility that security measures will fail provides no ‘proper ground’ for a broad-based attack on government information-collection practices.” *Nelson*, 131 S.Ct. at 763 (citing *Whalen*, 429 U.S. at 601). The Court emphasized that there is always a theoretical possibility of a data breach “any time the Government stores information.” *Nelson*, 131 S.Ct. at 763. *NASA*’s contract employees failed to establish that *NASA* uniquely exposed their personal information to heightened risk. *See id.* at 752 (implying that a holding for plaintiffs would result in an arbitrary “two-track” approach differentiating government contractors from government employees).

In sharp contrast, in this case the record management practices do routinely create an ongoing risk that patient data, which the Vermont law seeks to protect, will be disclosed. The risk of

reidentification of coupled with the opportunity to infer identity through the available data elements pose a substantial risk that information concerning sensitive medical conditions and prescription habits will be disclosed. These risks, established by *amici supra*, could have produced a different outcome in *Whalen*, and go far beyond what the Court called “mere possibility” in *NASA*.” *See Nelson*, 131 S.Ct. at 763.

At issue in this case is the effort to protect sensitive personal information from improper disclosure. There should be no doubt that the Court has recognized the importance of this interest. “Technology now permits millions of important and confidential conversations to occur through a vast system of electronic networks. These advances, however, raise significant privacy concerns. We are placed in the uncomfortable position of not knowing who might have access to our personal and business e-mails, our medical and financial records, or our cordless and cellular telephone conversations.” *Bartnicki*, 532 U.S. at 141, 151 (Rehnquist, C.J., dissenting). *See also Reno v. Condon*, 528 U.S. 141, 142 (2000) (The Drivers Privacy Protection Act “regulates the universe of entities that participate as suppliers to the market for motor vehicle information -- the States as initial suppliers of the information in interstate commerce and private resellers or redisclosers of that information in commerce.”)

Vermont has sought to protect the most sensitive of this private information from non-consensual disclosure to third parties. It is a sensible response to a serious problem.

**CONCLUSION**

*Amici* respectfully ask this Court to grant Petitioners' motion and reverse the decision of the Second Circuit and to remand to the district court with instructions to give full consideration to the privacy interests at issue in this matter.

Respectfully submitted,

MARC ROTENBERG  
JOHN VERDI  
SHARON GOOTT NISSIM  
THOMAS MOORE  
ELECTRONIC PRIVACY  
INFORMATION  
CENTER (EPIC)  
1718 Connecticut Ave. NW  
Suite 200  
Washington, DC 20009  
(202) 483-1140

March 1, 2011