Harvard Journal of Law & Technology
Volume 25, Number 1 Fall 2011

MUCH ADO ABOUT DATA OWNERSHIP

*Barbara J. Evans**

TABLE OF CONTENTS

I. INTRODUCTION 70

II. WHY DATA OWNERSHIP WOULD NOT PROTECT PATIENTS' PRIVACY 77
A. Nonconsensual Access to Patient Data Under a Property Regime 77
B. Nonconsensual Data Access Under the Existing Federal Regulations 82

III. WHY DATA OWNERSHIP CANNOT RESOLVE DATA ACCESS PROBLEMS 86
A. Identifying the Valuable Data Resources 90
1. Records of a Patient's Various Encounters with the Healthcare System 90
2. The Patient's Longitudinal Health Record ("LHR") 90
3. Longitudinal Population Health Data ("LPHD") 91
4. Unbiased LPHD 91
B. The Problem of Linking Data Across Healthcare Data Environments 93
C. Consent Bias and the Need for Nonconsensual Access to Patients' Health Data 95
D. The Role of Infrastructure and Demand-Side Factors 98
E. Why Data Propertization Proposals Fail 106

IV. THE HITECH ACT'S STRATEGY FOR PROMOTING INFRASTRUCTURE DEVELOPMENT 108
A. The Regulated Price of Infrastructure Services 108
B. Why the HITECH Act's Approach Offers Promise 110

V. WHAT STILL NEEDS TO BE DONE 113

---

* Professor; Co-director, Health Law & Policy Institute; Director, Center for Biotechnology & Law, University of Houston Law Center. Contact bjevans@central.uh.edu. J.D., Yale Law School; Ph.D., Stanford University; Post-doctoral Fellow, The University of Texas M.D. Anderson Cancer Center. This research was supported by the Greenwall Foundation and the University of Houston Law Foundation. The author would like to thank Baruch A. Brody, Robert A. Burt, Christine K. Cassel, Fred H. Cate, James W. Curran, Karla K.C. Holloway, Bernard Lo, Deven McGraw, Deborah Peel, Richard Platt, and Kristen Rosati for their insights and encouragement in this research.

A. Restoring the Proper Scope of the State's Police Power to Use Data to Promote Public Health 114
B. Developing a Workable Doctrine of Public Use of Private Data 119
1. How the Public Use Requirement Got Lost 120
2. Clarifying the Concept of Public Use of Data in Research 124
3. Developing a Workable Public Use Criterion 126
A. Focus Not on How Decisions Should Be Made, but by Whom 126
B. Develop Rules of Thumb for Identifying Suspect, "Non-Public" Uses of Data 127
C. Reject Utilitarian Balancing in Favor of Natural Rights Analysis 127

VI. CONCLUSION 129

I. INTRODUCTION

The Health Insurance Portability and Accountability Act of 1996 ("HIPAA")[1] Privacy Rule,[2] a major federal regulation affecting health information privacy, is criticized both for hindering access to health data[3] and for allowing too much data access.[4] Similar concerns[5] sur-

1. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended in scattered sections of 18, 26, 29 and 42 U.S.C.).

2. 45 C.F.R. pts. 160, 164 (2010).

3. COMM. ON HEALTH RESEARCH AND THE PRIVACY OF HEALTH INFO.: THE HIPAA PRIVACY RULE, INST. OF MED., BEYOND THE HIPAA PRIVACY RULE: ENHANCING PRIVACY, IMPROVING HEALTH THROUGH RESEARCH 66 (Sharyl J. Nass, Laura A. Levit & Lawrence O. Gostin eds., 2009) [hereinafter IOM, PRIVACY REPORT], *available at* http://www.nap.edu/catalog/12458.html; *see also* William Burman & Robert Daum, Infection Diseases Society of America, *Grinding to a Halt: The Effects of the Increasing Regulatory Burden on Research and Quality Improvement Efforts*, 49 CLINICAL INFECTIOUS DISEASES 328, 328 (2009) (arguing that "the application of the Health Insurance Portability and Accountability Act to research has overburdened institutional review boards (IRBs), confused prospective research participants, and slowed research and increased its cost"); Fred H. Cate, *Protecting Privacy in Health Research: The Limits of Individual Choice*, 98 CALIF. L. REV. 1765, 1797 (2010) ("Consent requirements [imposed by the HIPAA Privacy Rule] not only impede health research, but may actually undermine privacy interests.").

4. IOM, PRIVACY REPORT, *supra* note 3, at 66 (noting that the HIPAA Privacy Rule has not eliminated the concerns of the public, which is "deeply concerned about the privacy and security of personal health information," and reporting that "[i]n some surveys, the majority of respondents were not comfortable with their health information being provided for research except with notice and express consent").

5. *See* HHS, Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators, 76 Fed. Reg. 44,512, 44,523 (July 26, 2011) (to be codified at 45 C.F.R. pts. 46, 160, 164 and 21 C.F.R. pts. 50, 56) [hereinafter HHS, ANPRM] (noting in connection with a proposal to amend the Common Rule that "[c]ritics of the existing rules have observed that the current requirements for informed consent for future research with pre-existing data and biospecimens are

round the Federal Policy for the Protection of Human Subjects,[6] or "Common Rule."[7] While designed mainly to protect people who serve as subjects of experiments,[8] the Common Rule also applies to informational research[9] that merely uses people's data or biospecimens.[10] It thus plays an important role in regulating access to, and use of, people's health data. These regulations evolved over a twenty-eight-year period that began when the National Research Act of 1974[11] set up a National Commission[12] to develop the Common Rule. The period ended in 2002 when the HIPAA Privacy Rule was promulgated in its present form,[13] which draws heavily on concepts from the Common Rule.[14]

---

confusing and consume substantial amounts of researchers' and [Institutional Review Boards'] time and resources"); *id.* at 44,525 ("[O]ther fundamental protections for research participants may be warranted beyond updating the requirements for independent review and informed consent currently provided by the Common Rule.").

6. *See Federal Policy for the Protection of Human Subjects ("Common Rule")*, U.S. DEP'T OF HEALTH & HUMAN SERVS., http://www.hhs.gov/ohrp/humansubjects/commonrule/index.html (last visited Dec. 21, 2011).

7. 45 C.F.R. §§ 46.101–46.124 (2010).

8. *See* U.S. DEP'T OF HEALTH, EDUC., & WELFARE, PROTECTION OF HUMAN SUBJECTS: INSTITUTIONAL REVIEW BOARDS: REPORT AND RECOMMENDATIONS OF THE NATIONAL COMMISSION FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH, 43 Fed. Reg. 56,174 (Nov. 30, 1978) [hereinafter HEW, 1978 REPORT] (discussing the various types of research considered during development of the Common Rule); *see also infra* Part V (discussing the evolution of the Common Rule).

9. "Informational research" is one of many terms that refer to studies that use data and biospecimens. *See, e.g.,* HEW, 1978 REPORT, *supra* note 8, at 56,181–82 (using the older phrase "research using . . . records" to refer to informational research); BENGT D. FURBERG & CURT D. FURBERG, EVALUATING CLINICAL RESEARCH: ALL THAT GLITTERS IS NOT GOLD 29–37 (2d ed. 2007) (using the term "observational" to refer to methodologies that study data); David Casarett, Jason Karlawish, Elizabeth Andrews & Arthur Caplan, *Bioethical Issues in Pharmacoepidemiologic Research*, *in* PHARMACOEPIDEMIOLOGY 587, 588 (Brian L. Strom ed., 4th ed. 2005) (preferring the term "epidemiologic research"); IOM, PRIVACY REPORT, *supra* note 3, at 7 (distinguishing "information-based" research from clinical research); Brian L. Strom, *Study Designs Available for Pharmacoepidemiology Studies*, *in* PHARMACOEPIDEMIOLOGY, *supra* at 18, 21–26, (discussing the array of scientific methodologies — including observational studies that use existing data — for studying how people react to drugs).

10. 45 C.F.R. § 46.102(f) (2010) (defining "human subject" to include living individuals about whom an investigator obtains identifiable private information); *Guidance on Research Involving Coded Private Information or Biological Specimens*, U.S. DEP'T OF HEALTH & HUMAN SERVS., http://www.hhs.gov/ohrp/policy/cdebiol.html (last visited Dec. 21, 2011) [hereinafter *OHRP Guidance*] (interpreting how the Common Rule applies to research with data and biospecimens and describing circumstances in which such research will be subject to the Common Rule's informed consent requirements).

11. National Research Act of 1974 (National Research Service Award Act of 1974), Pub. L. No. 93-348, 88 Stat. 342 (codified as amended in scattered sections of 42 U.S.C.).

12. *See* HEW, 1978 REPORT, *supra* note 8, at 56,174 (discussing the National Commission and reporting its findings).

13. *See* Standards for Privacy of Individually Identifiable Health Information, 67 Fed. Reg. 53,182 (Aug. 14, 2002) (codified as amended at 45 C.F.R. pts. 160, 164) (implementing the currently effective version of the HIPAA Privacy Rule). *But see* Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the Health Information Technolo-

Regulatory approaches that have worked fairly well in clinical research[15] "are not easily exported and applied to the very different challenges of [informational] research."[16] As the Department of Health and Human Services ("HHS") was developing the HIPAA Privacy Rule in 2000, multiple public commenters, including several members of Congress, voiced this same concern.[17] Research with data and tissues has grown in importance,[18] making these problems more apparent and fueling calls for reform. In 2009, the Institute of Medicine ("IOM") called for changes to the HIPAA Privacy Rule.[19] The IOM recommended replacing the Privacy Rule with an unspecified "new approach" for regulating privacy and access to data for use in health research.[20] HHS recently published an advance notice of proposed rulemaking ("ANPRM"),[21] which called for changes to the Common Rule but offered few specifics, instead posing seventy-four broad questions for public comment.[22]

Rather than reforming the regulations, other proposals seek legislation to clarify data ownership.[23] Ownership of the data held in ad-

---

gy for Economic and Clinical Health Act, 75 Fed. Reg. 40,868 (July 14, 2010) (to be codified at 45 C.F.R. pts. 160, 164) (proposing changes to the HIPAA Privacy Rule that have not been finalized as of this writing).

14. *See* discussion *infra* Part II.B (discussing similarities between the Privacy Rule and the Common Rule).

15. *See* LAWRENCE M. FRIEDMAN, CURT D. FURBERG & DAVID L. DEMETS, FUNDAMENTALS OF CLINICAL TRIALS 2–5 (3d ed. 1998) (discussing clinical research, exemplified by a randomized, controlled clinical trial that monitors outcomes prospectively in two groups of people who either are or are not subjected to a particular treatment).

16. Casarett et al., *supra* note 9, at 587.

17. *See* Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,462, 82,690–91 (Dec. 28, 2000) (codified as amended at 45 C.F.R. pts. 160, 164) (responding to comments that it was inappropriate for the HIPAA waiver provision to be "modeled on the existing system of human subject protections" and that "the Common Rule's requirements may be suited for interventional research involving human subjects, but is ill suited to the archival and health services research typically performed using medical records without authorization").

18. *See, e.g.*, Fred D. Brenneman et al., *Outcomes Research in Surgery*, 23 WORLD J. SURGERY 1220 (1999) (discussing the growth of informational research after 1980); *Outcomes Research Fact Sheet (2000)*, AGENCY FOR HEALTHCARE RESEARCH AND QUALITY, http://www.ahrq.gov/clinic/outfact.htm (last visited Dec. 21, 2011) [hereinafter AHRQ, *Fact Sheet*] (same); *see also* Barbara J. Evans & Eric M. Meslin, *Encouraging Translational Research Through Harmonization of FDA and Common Rule Informed Consent Requirements for Research with Banked Specimens*, 27 J. LEGAL MED. 119, 122 (2006) (discussing the growing importance of research with biospecimens); Rina Hakimian & David Korn, *Ownership and Use of Tissue Specimens for Research*, 292 JAMA 2500, 2500 (2004) (same).

19. IOM, PRIVACY REPORT, *supra* note 3, at 28–29 (summarizing the IOM's recommendations).

20. *Id.* at 28.

21. *See* HHS, ANPRM, *supra* note 5, at 44,512.

22. *Id.* at 44,517–29.

23. *See, e.g.*, Mark A. Hall*, Property, Privacy, and the Pursuit of Interconnected Electronic Medical Records*, 95 IOWA L. REV. 631, 651 (2010) ("[I]f patients were given owner-

ministrative[24] and clinical databases is a matter of state law and, in most states, data ownership is not clearly defined.[25] Patient data ownership is touted by some observers as a way to enhance patient privacy[26] and by others as a way to make data more widely available for

---

ship of their complete medical treatment and health histories, they could license to compilers their rights to that information in a propertized form that could be more fully developed and commercialized."); Marc A. Rodwin, *Patient Data: Property, Privacy & the Public Interest*, 36 AM. J.L. & MED. 586, 589 (2010) (arguing for public ownership of de-identified patient data).

24. *See* Leslie L. Roos et al., Inst. of Med., *Strengths and Weaknesses of Health Insurance Data Systems for Assessing Outcomes*, *in* MODERN METHODS OF CLINICAL INVESTIGATION 47, 47, 52 (Annetine C. Gelijns ed., 1990), *available at* http://www.nap.edu/openbook.php?record_id=1550 [hereinafter IOM, MODERN METHODS] (discussing health research that uses administrative data, such as claims data held by Medicare, Medicaid, and private health insurers).

25. *See* Barbara J. Evans, *Congress' New Infrastructural Model of Medical Privacy*, 84 NOTRE DAME L. REV. 585, 596–97 (2009) (discussing the current status of ownership of health data and human tissue specimens); David L. Silverman, *Data Security Breaches: The State of Notification Laws*, 19 INTELL. PROP. & TECH. L.J., July 2007, at 5, 8 (discussing the "precarious" nature of ownership of database content under current law). *But see* Susan E. Gindin, *Lost and Found in Cyberspace: Informational Privacy in the Age of the Internet*, 34 SAN DIEGO L. REV. 1153, 1195 n.231 (1997) (noting several cases where courts have recognized patients' ownership of medical records); Seth Axelrad, *State Statutes Declaring Genetic Information to be Personal Property*, AM. SOC'Y OF L., MED. & ETHICS, http://www.aslme.org/dna_04/reports/axelrad4.pdf (last visited Dec. 21, 2011) (listing four states — Alaska, Colorado, Florida, and Georgia — that recognize individual property rights in one category of health data, genetic information).

26. *See, e.g.*, Deborah C. Peel, *Written Testimony Before the HIT Policy Committee*, ELECTRONIC PRIVACY INFO. CENTER (Sept. 18, 2009), http://epic.org/privacy/medical/Peel_PPR%20Written%20testimony%20HIT%20Policy%20Committee.pdf (indicating that "twenty focus groups held across the country in order to understand consumers' awareness, beliefs, and fears concerning [health information technology]" discovered that "[a] majority want to 'own' their health data"); *Principles: More Patient Privacy Principles*, PATIENT PRIVACY RIGHTS, http://patientprivacyrights.org/principles (last visited Dec. 21, 2011) (listing a set of eleven "Patient Privacy Principles [that] should be included in all Health [information technology] legislation" and listing as the first such principle, "[r]ecognize that patients own their health data"). Patient data ownership advocates often ground their position in the claim that privacy is a state in which patients exercise full control over their own health information. *See* Peel, *supra* (framing privacy as "control of personal information" and "consumer control over [personal health information]"). Patients want assured access to the data about themselves that are maintained or stored by others. *See* Richard H. Thaler, *Show Us the Data. (It's Ours, After All.)*, N.Y. TIMES, Apr. 23, 2011, at BU4 (discussing "a host of privacy issues" raised by collection and dissemination of personal data and calling for persons whose data is stored to have access to it); Elizabeth Cohen, *Patients Demand: "Give Us Our Damned Data"*, CNN (Jan. 14, 2010), http://articles.cnn.com/2010-01-14/health/medical.records_1_hospital-bed-patients-demand-medical-records; Leslie A. Saxon, *Owning Your Health Information: An Inalienable Right*, THE HUFFINGTON POST (Oct. 7, 2009), http://www.huffingtonpost.com/leslie-a-saxon-md/owning-your-health-inform_b_312852.html (discussing data ownership as a way to enhance patients' access to data). Another desired aspect of privacy is for patients to be able to control others' access to and use of the patients' data). *See A Declaration of Health Data Rights*, HEALTHDATARIGHTS.ORG, http://www.healthdatarights.org (last visited Dec. 21, 2011) (advocating a "self-evident and inalienable" entitlement not only to "[h]ave the right to our own health data" but also to "have the right to share our health data with others as we see fit").

research.[27] Still others call for public (governmental) ownership to enhance researchers' access to data.[28] While differing in details, data propertization proposals seem to agree that property rights in data are important and that clarifying them should be high on the legislative agenda.[29] Ominously, this view is starting to infect policymakers,[30] raising a real risk that what began as an abstract scholarly debate may end in ill-advised legislation.

The urge to propertize health data needs to be weighed skeptically and with a clear understanding of how property rights actually work. If pursued, data ownership may disappoint many of its proponents because of a surprising truth: the framework of patient entitlements and protections afforded by the HIPAA Privacy Rule and the Common Rule is strikingly similar to what patients would enjoy if they owned their data.[31] Part II challenges the claim that private data ownership would improve privacy protection. It finds that both regimes — patient ownership of data, on the one hand, and the federal regulatory protections, on the other — provide pliability-rule protection[32] that strikes a balance between patient control and the public's need for data access. Both regimes allow some unconsented uses of patients'

---

27. *See* Hall, *supra* note 23, at 651; *see also* Mark A. Hall & Kevin A. Schulman, *Ownership of Medical Information*, 301 JAMA 1282, 1283–84 (2009) (discussing advantages of patient-controlled longitudinal health records and suggesting that one way to foster the development of such records would be to "give patients the rights to sell access to their records, rights that are superior to the property rights held by [entities that currently hold patients' data]").

28. *See* Marc A. Rodwin, *The Case for Public Ownership of Patient Data,* 302 JAMA 86, 87–88 (2009) (arguing for governmental ownership of de-identified patient data).

29. *See, e.g.*, Hall, *supra* note 23, at 637 ("The law's uncertainty over ownership and control of medical information is widely regarded as a major barrier to effective networking of [electronic medical records], and policy analysts consider the legal status of medical information to be a critical question at or near the top of issues needing resolution."); *id.* at 631 ("How this issue is resolved can determine how or whether massive anticipated developments in electronic health records will take shape."); Rodwin, *supra* note 23, at 586 ("How the law defines ownership of patient data will shape whether its benefits can be developed and also affects patient confidentiality.").

30. *See, e.g.*, H. COMM. ON PUB. HEALTH, INTERIM REP. TO THE 82ND TEX. LEG., H. 82-C410, 1st Sess., at 17, 20–21 (2010) (stating as its first recommendation that "[t]he Legislature should determine clearly in law who is the owner of medical records").

31. *See* discussion *infra* Part II; *see also* Hall & Schulman, *supra* note 27, at 1282 (acknowledging that "the effect of other legal regimes may sometimes resemble property law").

32. *See* Abraham Bell & Gideon Parchomovsky, *Pliability Rules*, 101 MICH. L. REV. 1 (2002). The important feature of pliability rule protection, for purposes of this discussion, is that it offers a dynamic scheme of entitlements in which a baseline rule of consensual ordering of data access can shift to nonconsensual access under specified circumstances. *See id.* at 5 ("Pliability, or pliable, rules are contingent rules that provide an entitlement owner with property rule or liability rule protection as long as some specified condition obtains; however, once the relevant condition changes, a different rule protects the entitlement — either liability or property, as the circumstances dictate. Pliability rules, in other words, are dynamic rules, while property and liability rules are static.").

data, and the grounds for nonconsensual data use are substantively similar under either regime. This similarity suggests that property rights may not be the right locus for reform. Creating property rights in data would produce a new scheme of entitlements that is substantively similar to what already exists, thus perpetuating the same frustrations all sides have felt with the existing federal regulations.

Part III challenges the claim that clarifying data ownership would improve access to useful data resources for clinical care, public health,[33] and research. This claim was central to the recent debate between Professors Hall and Schulman[34] and Professor Rodwin,[35] who disagreed whether private or public ownership of patients' data would better promote data access. Data propertization proposals fail because patients' raw health information is not in itself a valuable data resource, in the sense of being able to support useful, new applications.[36] Creating useful data resources requires significant inputs of human and infrastructure services, and owning data is fruitless unless there is a way to acquire the necessary services.[37] Part IV considers the impact of the 2009 Health Information Technology for Economic and Clinical Health ("HITECH") Act,[38] which authorized data holders[39] to supply the needed services commercially, subject to regulated

33. Public health data uses include tracking the spread of communicable diseases, reporting injuries that suggest child abuse, and conducting postmarket surveillance to monitor the safety of approved drugs. *See* LAWRENCE O. GOSTIN, PUBLIC HEALTH LAW 4 (2d ed. 2008) (characterizing public health law as focusing on population-oriented (as opposed to patient-specific) efforts "to ensure the conditions for people to be healthy (to identify, prevent, and ameliorate risks to health in the population)" and "to pursue the highest possible level of physical and mental health in the population, consistent with the values of social justice"); Paul J. Amoroso & John P. Middaugh, *Research vs. Public Health Practice: When Does a Study Require IRB Review?*, 36 PREVENTIVE MED. 250, 250 (2003) (providing as examples of public health activities mandatory reporting of communicable diseases, investigating disease outbreaks, and collecting confidential information by public health authorities to protect the public health); Evans, *supra* note 25, at 589, 614–18 (discussing postmarket drug safety surveillance and its character as a public health activity).

34. *See* Hall, *supra* note 23; Hall & Schulman, *supra* note 27.

35. *See* Rodwin, *supra* note 23; Rodwin, *supra* note 28.

36. *See* discussion *infra* Parts III.A–III.C.

37. *See* discussion *infra* Part III.D.

38. 42 U.S.C.A. §§ 17931–17940 (West 2010 & Supp. 2011).

39. "Data holder" is one of many names for entities — such as physicians, hospitals, insurers, and other health database operators — that possess individuals' health data. *See FDA's Sentinel Initiative*, U.S. FOOD & DRUG ADMIN., http://www.fda.gov/Safety/FDAsSentinelInitiative/ucm2007250.htm (last visited Dec. 21, 2011). *But see* JANET M. MARCHIBRODA, EHEALTH INITIATIVE FOUND., DEVELOPING A GOVERNANCE AND OPERATIONS STRUCTURE FOR THE SENTINEL INITIATIVE 21, 34 (2009), *available at* http://www.regulations.gov/#!documentDetail;D=FDA-2009-N-0192-0006 (using "data environment" and "data source"); Meryl Bloomrosen & Don Detmer, *Advancing the Framework: Use of Health Data — A Report of a Working Conference of the American Medical Informatics Association*, 15 J. AM. MED. INFO. ASS'N. 715, 715 (2008) (preferring "data steward").

pricing.[40] This approach, which draws on a long tradition of successful American infrastructure regulation, offers promise in resolving infrastructure bottlenecks, which — rather than the unresolved status of data ownership — have presented the key impediment to data availability.[41]

Despite this progress, important problems remain unresolved. A major challenge in twenty-first century privacy law and research ethics will be to come to terms with the inherently collective nature of knowledge generation in a world where large-scale informational research is set to play a more prominent role.[42] Informational research differs starkly from interventional research, exemplified by randomized, controlled clinical trials,[43] which were the major workhorse of late twentieth-century biomedical discovery.[44] A person's refusal to participate in a clinical trial does not jeopardize the broader clinical research enterprise, which can move forward using other willing research subjects; only 600–3000 people are needed for a typical clinical drug trial.[45] In contrast, a person's refusal to participate in informational research may bias the dataset and reduce its statistical power for everyone.[46] Many important types of informational research must be done collectively with large, inclusive datasets.[47] An individual's wish not to participate, perhaps motivated by privacy concerns, potentially places other human beings at risk and undermines broader public interests — for example, in public health or medical discovery — in which the individual shares.[48] Existing regulations lack tools to resolve this complex dilemma.

---

40. *See* discussion *infra* Part IV.

41. *See* discussion *infra* Parts III and IV.

42. *See* 21 U.S.C.A. § 355(o)(3)(D) (West. 2010 & Supp. 2011) (placing the U.S. Food and Drug Administration under a requirement to consider and reject the use of observational studies before the agency can order a postmarket clinical drug trial); ROUNDTABLE ON EVIDENCE-BASED MED., INST. OF MED., THE LEARNING HEALTHCARE SYSTEM: WORKSHOP SUMMARY 128, 130 (LeighAnne Olsen et al. eds., 2007) [hereinafter IOM, LEARNING HEALTHCARE] (discussing the growing use of observational methodologies); Barbara J. Evans, *Seven Pillars of a New Evidentiary Paradigm: The Food, Drug, and Cosmetic Act Enters the Genomic Era*, 85 NOTRE DAME L. REV. 419, 479–85 (2010) (same).

43. *See* FRIEDMAN ET AL., *supra* note 15, at 2 (noting that clinical trials involve the use of intervention techniques); FURBERG & FURBERG, *supra* note 9, at 11–22 (discussing randomized, controlled clinical trials).

44. *See* Evans, *supra* note 42, at 432 (noting that randomized, controlled clinical trials played a central role in the mid-twentieth-century drug approval paradigm that the FDA implemented under the 1962 Drug Amendments).

45. COMM. ON THE ASSESSMENT OF THE U.S. DRUG SAFETY SYS., INST. OF MED., THE FUTURE OF DRUG SAFETY 36 (Alina Baciu et al. eds., 2007), *available at* http://books.nap.edu/openbook.php?record_id=11750.

46. *See* discussion *infra* Part III.C.

47. *Id.*

48. *See* Paul Starr, *Health and the Right to Privacy*, 25 AM. J.L. & MED. 193, 200 (1999) (noting that patients not only have an interest in the privacy of their health records, but also have an interest in research and other efforts to improve the medical care available to them).

Part V identifies two fundamental defects of the current HIPAA Privacy Rule and Common Rule. First, these regulations conceive the state's police power to use data to promote public health much more narrowly than the police power is conceived in all other legal contexts. This has the effect of thwarting legitimate uses of data to improve the public's health. Second, the existing provisions for approving nonconsensual research uses of data fail to incorporate any public use requirement — that is, a requirement that unconsented data uses must be justified by a publicly beneficial purpose. As things stand, persons whose health data are used in research have no assurance that the use will serve any socially beneficial purpose at all.[49] The 2009 IOM recommendations,[50] the more recent ANPRM,[51] and the ongoing clamor of data propertization proposals all fail to address these problems. Reaping the full benefit of interoperable health information systems requires coming to grips with them. It is time to reframe the debate. The right question is not who owns health data, nor is it any of the seventy-four queries that erupted from the recent ANPRM. Instead, the debate should be about appropriate public uses of private data and how best to facilitate these uses while adequately protecting individuals' interests.

II. WHY DATA OWNERSHIP WOULD NOT PROTECT PATIENTS' PRIVACY

A. Nonconsensual Access to Patient Data Under a Property Regime

Data propertization proposals fall into two broad categories: pro-privacy proposals that portray private ownership as a way to bolster patients' power to block unwanted uses of their data and pro-access proposals that aim to promote wider availability of data for clinical, research, and public health uses.[52] The pro-privacy proposals rest on a mythical view of private property. Three centuries ago, Sir William Blackstone noted how the human imagination is drawn to the idea of property as "that sole and despotic dominion which one man claims and exercises over external things of the world in total exclusion of

49. *See* discussion *infra* Part V.B (explaining the complete absence of a public use requirement in the Common Rule and HIPAA waiver provisions, which allow nonconsensual access to data for research).

50. IOM, PRIVACY REPORT, *supra* note 3, at 28–29 (summarizing the IOM's recommendations).

51. HHS, ANPRM, *supra* note 5.

52. *See supra* notes 23, 26–28 and accompanying text (providing examples of various pro-privacy and pro-access proposals).

the right of any other individual in the universe."[53] This idea resonates with the "autonomy *über alles*" strand of privacy advocacy that asserts that a patient's right to control access to health data should trump all other interests, even society's interest in conducting studies that might save or improve other people's lives.[54] Blackstone, however, was merely describing how people *imagine* property. He himself did not espouse this view,[55] nor has American law ever done so.[56]

Different assets call for different forms of ownership, and proponents of patient data ownership do not always specify what they have in mind.[57] Data ownership might, for example, need to look something like the nonexclusive rights riparian owners have in a river that runs by their land — a right to use the river oneself but not to interfere with others' simultaneous uses for fishing and navigation[58] — or like a copyright, which expires after a fixed term of years and allows fair use by others even during that term.[59] Pro-privacy proposals seem to draw on the ideal of property reflected in the saying, "one's home is one's castle."[60] In the usual course of events, access to a person's home requires a consensual transaction with the owner, and unconsented uses can be enjoined.[61] This package of rights and remedies

---

53. 2 WILLIAM BLACKSTONE, COMMENTARIES *2, *available at* http://avalon.law.yale.edu/18th_century/blackstone_bk2ch1.asp (spelling conformed to modern conventions).

54. *See* Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,462, 82,698 (Dec. 28, 2000) (codified as amended at 45 C.F.R. pts. 160, 164) (noting in the preamble to the HIPAA Privacy Rule that some public comments had opposed waiver provisions that would have let data be used without consent when "the research is of sufficient importance so as to outweigh the intrusion of the privacy of the individual whose information is subject to the disclosure," and providing an example of one pro-autonomy commenter who insisted that "common purposes should not override individual rights in a democratic society").

55. 1 BLACKSTONE, *supra* note 53, at *119, *available at* http://avalon.law.yale.edu/18th_century/blackstone_bk1ch1.asp (recognizing that rights include both rights that are owed to individuals and those that are "due from every citizen, which are usually called civil duties" (spelling conformed to modern conventions)).

56. *See* Eric R. Claeys, Kelo*, the Castle, and Natural Property Rights*, *in* PRIVATE PROPERTY, COMMUNITY DEVELOPMENT, AND EMINENT DOMAIN 35, 40–43 (Robin Paul Malloy ed., 2008) (discussing the natural rights theory reflected in eighteenth- and nineteenth-century American takings jurisprudence and how it allowed interference with property rights under certain circumstances).

57. *See, e.g.*, Hall, *supra* note 23, at 663 (calling for an unspecified "right mix and forms of property rights among patients, providers, researchers, and compilers").

58. *See* Eric R. Claeys, *Takings, Regulations, and Natural Property Rights*, 88 CORNELL L. REV. 1549, 1591–93, 1596, 1600–02 (2003**)** (discussing cases that have addressed conflicts between riparian owners' rights to build dams, mills, wharves, and other structures and the need to preserve other uses such as navigation and fishing).

59. Abraham Bell, *Private Takings*, 76 U. CHI. L. REV. 517, 541 (2009).

60. Claeys, *supra* note 56, at 35–36 (discussing the popular meaning of the castle metaphor).

61. Thomas W. Merrill, *The Economics of Public Use*, 72 CORNELL L. REV. 61, 64 (1986) (discussing rights and remedies available under a scheme of property-rule protection).

corresponds to property-rule protection,[62] and it is what privacy proponents seem to be seeking in their calls for data ownership: consensual ordering of data access and the power to stop unconsented uses.

The fatal flaw in pro-privacy proposals is this: having a property right does not ensure property-rule protection. Law recognizes that there are many situations where consensual transactions cannot be relied on as a way of ordering an owner's relations with the larger community.[63] In many circumstances, a property owner only receives liability-rule protection,[64] which means the owner can be forced to give up her property in return for compensation that is externally set, often by a court, legislature, or administrative agency.[65] That compensation may be zero. The government — when acting under its police power to protect the public's health, safety, morals, or welfare — has broad power to confiscate or interfere with property without compensating the owner.[66] Dating back to colonial times, the state's police power has been used not just to prevent property owners from injuring others, but also to pursue broader public welfare objectives for the benefit of the community.[67] "[T]here was no single paradigm of public welfare that confined what we now call the police power. Then, as now, lawmakers pursued a shifting amalgam of goals . . . . Legislation coercively promoted uses of private land that were viewed as conducive to the community's well-being."[68] Consistent with this tradition, the government can require nonconsensual access to data for use in public health activities, which long have been viewed as a legitimate exercise of the state's police power.[69] This would remain true even if data were patient-owned.

---

62. Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972).

63. *See id.* at 1108–09 (discussing the problems with relying on consensual transactions to compensate for accidents); Bell & Parchomovsky, *supra* note 32, at 8–19 (discussing the evolution of entitlement theory as it bears on the relative merits of consensual and nonconsensual ordering in various circumstances).

64. Calabresi & Melamed, *supra* note 62, at 1092.

65. Bell & Parchomovsky, *supra* note 32, at 3.

66. *See* Merrill, *supra* note 61, at 66 ("[T]he government [can] take [a citizen's] property without his consent and without compensation . . . when the government legitimately exercises its power to tax or its police power.").

67. *See, e.g.*, John F. Hart, *Land Use Law in the Early Republic and the Original Meaning of the Takings Clause*, 94 NW. U. L. REV. 1099, 1102 (2000); *id.* at 1107 (discussing historical uses of the state's police power to require owners to confer positive externalities on the community); William Michael Treanor, *The Original Understanding of the Takings Clause and the Political Process*, 95 COLUM. L. REV. 782, 797 (1995) (same).

68. Hart, *supra* note 67, at 1107.

69. Hall, *supra* note 23, at 659 (noting that the government currently requires disclosure of identifiable information for public health purposes under its police power); *see also* GOSTIN, *supra* note 33, at 11 ("Public health has historically constrained the rights of individuals and businesses so as to protect community interests in health . . . [including] the use of reporting requirements affecting privacy . . . ."); Wendy E. Parmet, *After September 11:*

The state also has eminent domain power to take property for "public use"[70] without the owner's consent, subject to payment of just compensation.[71] The public uses that can support a taking are quite broad and could include private, commercial research uses of data, if data were patient-owned. Takings require "some showing of 'publicness'" of the intended use,[72] and takings that lack the requisite public quality can be enjoined.[73] Public uses traditionally involved placing the property under public ownership or transferring it to a private company, such as a utility or railroad, that is obligated to serve the public, often but not always for a regulated price.[74] There was never a requirement that the fruits of a taking be made *freely* available to the public: railroads and stadiums built on taken land routinely require users to buy tickets.[75] Modern courts, somewhat controversially,[76] allow takings that transfer property to new private owners for commercial projects that need not be open to the general public[77] and for projects that offer only indirect public benefits, such as boosting local tax revenues or aiding urban renewal or land reform.[78]

The possibility of eminent domain appears to have been lost on privacy advocates who view data ownership as a way to halt unconsented, private-sector research use of data. Modern takings doctrine

---

Rethinking Public Health Federalism, 30 J.L. MED. & ETHICS 201, 202–03 (2002) (discussing the police power to protect public health).

70. *See* Robin Paul Malloy & James Charles Smith, *Private Property, Community Development, and Eminent Domain*, *in* PRIVATE PROPERTY, COMMUNITY DEVELOPMENT, AND EMINENT DOMAIN, *supra* note 56, at 1, 8 (discussing the public use requirement).

71. U.S. CONST. amend. V.

72. Merrill, *supra* note 61, at 61.

73. *Id.* at 68, 85.

74. *See* Kelo v. City of New London, 545 U.S. 469, 479 (2005) ("[M]any state courts in the mid-nineteenth-century endorsed 'use by the public' as the proper definition of public use . . . ."); *see also* Claeys, *supra* note 56, at 42–43 (discussing early cases that applied natural rights theory to limit the use of eminent domain to situations where the taken property would go "to a state agency, or to a private entity with standard common-carrier duties of non-discrimination and reasonable rates"); Bell, *supra* note 59, at 552 ("[P]ublic ownership of the taken property is not a necessary companion to just and efficient takings . . . .").

75. *See* Brett M. Frischmann, *An Economic Theory of Infrastructure and Commons Management*, 89 MINN. L. REV. 917, 925–26 (2005).

76. *See* Michael Allan Wolf, *Hysteria Versus History: Public Use in the Public Eye*, *in* PRIVATE PROPERTY, COMMUNITY DEVELOPMENT, AND EMINENT DOMAIN, *supra* note 56, at 15, 15–20 (discussing public reaction to cases that have relied on a broader public purpose test).

77. *See, e.g.*, *Kelo*, 545 U.S. at 472–75, 484 (allowing homes to be transferred to private corporations for use in constructing an office park that, when completed, would be available to commercial tenants rather than to the general public); Poletown Neighborhood Council v. City of Detroit, 304 N.W.2d 455, 459–60 (Mich. 1981) (allowing property to be taken for development of a private manufacturing plant).

78. *See, e.g.*, *Kelo*, 545 U.S. at 472 (allowing taking for purpose of generating tax revenue); Haw. Hous. Auth. v. Midkiff, 467 U.S. 229 (1984) (allowing taking for purpose of land reform); Berman v. Parker, 348 U.S. 26 (1954) (allowing taking for purpose of urban renewal).

would allow privately owned health data to be taken for use in academic and commercial research that offers a prospect of developing a beneficial therapy. This is true even if the new therapy, when successfully developed, would be available only to patients who can pay for it. It seems doubtful that patients would be entitled to compensation when their data were taken for use in research. Courts construe "just compensation" to mean payment of market value[79] — what the property would fetch in an alternative, consensual sale on the open market.[80] There is no compensation for subjective value, such as the emotional attachment an owner has to a particular home, or for undeveloped use rights — what the undeveloped property might have been worth if the current owner had chosen to develop it.[81] There also is no compensation for consequential costs of the taking, such as an owner's moving expenses.[82] These same limitations presumably would apply if patient-owned data were taken for public use in research. When patients wish to have their data "lie fallow" because of privacy concerns, the fair market value of the data arguably is zero: if patients oppose having their data used in research at all, there is no alternative consensual use by which to gauge the data's market value. The value of unused data is largely subjective, reflecting an emotional attach-

79. *See* Abraham Bell & Gideon Parchomovsky, *The Hidden Function of Takings Compensation*, 96 VA. L. REV. 1673, 1677 (2010) (noting that eminent domain only compensates market value, not emotional attachment or subjective valuation).

80. *See* Merrill, *supra* note 61, at 83 (noting that compensation is assessed relative to the property's fair market value "in its highest and best use *other than* the use proposed by the condemnor").

81. *Id.*; *see also* Claeys, *supra* note 58, at 1646–47 (noting that takings compensation is based on adverse economic impact in the form of interference with "distinct investment-backed expectations" (citing Penn Central Transp. Co. v. City of New York, 438 U.S. 104, 124 (1978))).

82. *See* Merrill, *supra* note 61, at 83 ("[O]ther personal losses which do not 'run' with the property, such as lost goodwill, consequential damages to other property, relocation costs, and attorney fees, are also not compensable."); Frank I. Michelman, *Property, Utility, and Fairness: Comments on the Ethical Foundations of "Just Compensation" Law*, 80 HARV. L. REV. 1165, 1254 (1967) ("When land is appropriated for clearance and redevelopment, its owner is, of course, compensated in the amount of its 'fair market value.' But, by the generally received doctrines, . . . tenant-owners are not constitutionally entitled to be compensated for the disruptive effects of changing neighborhoods and sinking new roots, or even, in case a business is uprooted, for good will destroyed, or, very possibly, for the cash outlay entailed in moving."). *But see* Malloy & Smith, *supra* note 70, at 8 (noting that some jurisdictions may adjust compensation in light of relocation expenses and costs to acquire substitute property).

ment to the data and a wish to keep it secret.[83] This is not compensable under modern takings doctrine.[84]

B. Nonconsensual Data Access Under the Existing Federal Regulations

The HIPAA Privacy Rule and the Common Rule offer a framework of patient entitlements and protections that is strikingly similar to what patients would enjoy if they owned their data. Under ordinary circumstances, both regulations require consensual ordering of data access: they require a privacy authorization[85] or informed consent[86] before data can be used. However, both regulations contain exemptions, exceptions, and definitional nuances that shift to a regime of liability-rule protection under certain circumstances.[87] The recent ANPRM would adjust specific provisions but preserve this same overall structure.[88]

Under the current regulations, certain activities that are considered to have high social value — such as using data for judicial, law enforcement, and public health purposes — are not subject to the usual consent and authorization requirements.[89] Nonconsensual research

---

83. *See, e.g.*, ALAN F. WESTIN, HOW THE PUBLIC VIEWS PRIVACY AND HEALTH RESEARCH 13 (2007) (suggesting that the fifty percent of surveyed individuals who expressed concern about research access to their health data "are expressing the 'pure privacy' position — a sense of violation or intrusion if their sensitive health information is seen by an unknown third party, . . . even if a promise of anonymity is offered; and even if no actual harm to reputation is likely to result from such research activity").

84. *See* Hall, *supra* note 23, at 659; Rodwin, *supra* note 23, at 609. Hall and Rodwin both argue that patients have no property interest in de-identified data; therefore, research uses of such data would not constitute a taking and, hence no compensation would be owed. My point is different: even if the data were identifiable or fully identified, and even if the patient had a property interest in the data, and even if research use of the data were deemed to be a taking, it is still true that no compensation would be owed, because takings doctrine does not compensate subjective value — and the perceived value of keeping data unutilized is a subjective value.

85. *See* 45 C.F.R. § 164.508 (2010) (describing authorization requirements of the HIPAA Privacy Rule).

86. *Id.* § 46.116 (describing informed consent requirements of the Common Rule).

87. *See id.* § 164.512 (describing exceptions to the HIPAA Privacy Rule's authorization requirements); *see also id.* § 46.101(b)–(d) (describing exemptions to the Common Rule); *id.* § 46.102(d), 46.102(f) (defining the terms "research" and "human subject"); *OHRP Guidance*, *supra* note 10 ("OHRP does not consider research involving *only* coded private information or specimens to involve human subjects as defined under 45 C.F.R. § 46.102(f) . . . .").

88. *See, e.g.*, HHS, ANPRM, *supra* note 5, at 44,518–21, 44,527 tbl.1 (proposing to require consent for certain exempt data uses that do not presently require consent under the Common Rule). *But see id.* at 44,523–25 (contemplating circumstances under which consent requirements could still be waived); *id.* question 24, at 44,521 (seeking public comment on whether certain high-valued activities, such as quality improvement and public health activities, should lie outside the Common Rule's consent requirements).

89. *See* Barbara J. Evans, *Issue Brief: Appropriate Human-Subject Protections for Research Use of Sentinel System Data*, FDA SENTINEL INITIATIVE MEETING SERIES: LEGAL

use of data is allowed under conditions aimed at reducing privacy risks to the data subjects. Such use is allowed if the data have been de-identified,[90] coded in compliance with specified standards,[91] or converted to a limited data set.[92] Nonconsensual research uses are also allowed if an Institutional Review Board or privacy board (collectively, "IRB")[93] approves a waiver of the usual consent or authorization requirements.[94] Data supplied to researchers under a HIPAA waiver must meet "minimum necessary"[95] requirements — i.e., no more information can be disclosed than is necessary to accomplish the intended research purpose. However, there is no requirement that the data be de-identified or even coded to qualify for a waiver. In theory, it is possible to disclose fully identified data under a waiver, if the

---

ISSUES IN ACTIVE MEDICAL PRODUCT SURVEILLANCE 4 (Mar. 2010), *available at* http://www.brookings.edu/~/media/Files/events/2010/0308_FDA_legal_issues/Panel%203%20Issue%20Brief.pdf (summarizing the various pathways for nonconsensual use of data under the HIPAA Privacy Rule and the Common Rule); Evans, *supra* note 25, at 597, 619–22 (describing the provisions for nonconsensual data access under the HIPAA Privacy Rule); *id.* at 625–30 (describing nonconsensual access to data under the Common Rule and under the FDA's human subject protection regulations); *see also* Kristen Rosati, *An Analysis of Legal Issues Related to Structuring FDA Sentinel Initiative Activities*, U.S. FOOD & DRUG ADMIN. (2009), *available at* http://www.regulations.gov/#!documentDetail;D=FDA-2009-N-0192-0003 (providing a detailed examination of provisions of the Privacy Rule, Common Rule, the Privacy Act, and other relevant laws — such as those governing data on substance abuse — that affect access to data used in FDA's postmarket drug safety surveillance activities).

90. *See* 45 C.F.R. § 164.514(b) (2010) (allowing data to be de-identified, for purposes of HIPAA, by removing eighteen specific types of identifiers or by having a statistician certify that the risk of re-identification is "very small"); *id.* § 46.102(f) (defining "human subject" in a way that means that research with data is not covered by the Common Rule's consent requirements if investigators do not receive identifying information or interact with the subjects). *But see* HHS, ANRPM, *supra* note 5, at 44,519, 44,523 (requiring consent for some uses of de-identified data that would not require consent under the existing Common Rule).

91. *See* 45 C.F.R. § 164.514(c) (2010) (allowing coded data to be considered "de-identified" under the HIPAA Privacy Rule if the code key is subject to certain restrictions on derivation and access); *OHRP Guidance*, *supra* note 10 (discussing permissible coding arrangements under the Common Rule).

92. *Id.* § 164.514(e).

93. *See* 45 C.F.R. §§ 46.103(b), 46.107–108 (2010) (describing IRBs: private ethical review panels, often staffed by employees of the data holder or data-using research institution, to which the Common Rule delegates various aspects of research oversight); *id.* § 164.512(i)(2)(iv) (allowing of waivers of consent under the HIPAA Privacy Rule to be approved by either a Common Rule-compliant IRB or by a HIPAA-compliant "privacy panel" that is similar to an IRB); *see also* Evans, *supra* note 25, at 622–25 (discussing and critiquing the role of IRBs in approving consent waivers).

94. 45 C.F.R. § 164.512(i) (2010) (HIPAA waiver provision); *id.* § 46.116(d) (Common Rule waiver provision).

95. 45 C.F.R. § 164.514(d) (2010).

research requires the use of identified data and if an IRB deems the other waiver conditions to be met.[96]

While some people object to any nonconsensual use of their data, there is fairly solid public support for police power uses of data — such as monitoring the spread of epidemics — that protect public health, safety, and welfare.[97] The public also has some degree of comfort with the use of de-identified and other "masked" forms of data[98] despite ongoing concerns about the potential for such data to be re-identified.[99] Waivers do not inspire similar levels of public understanding.[100] They are subject to ongoing critique from research institutions and IRBs that find the waiver provisions cumbersome to apply[101] and from scholars and privacy advocates who view them as an abuse-prone bypass to consent requirements.[102]

The waiver provisions of the HIPAA Privacy Rule and the Common Rule are best understood as a regulator-created analogue of private takings power. These provisions let private bodies (IRBs) approve nonconsensual research use of data.[103] There is a long history in the United States, dating back to colonial times, of delegating takings power to private parties — such as developers of milldams and railroads — so that they can take property directly for socially benefi-

96. *See* Barbara J. Evans, *Ethical and Privacy Issues in Pharmacogenomic Research*, *in* PHARMACOGENOMICS: APPLICATIONS TO PATIENT CARE 313, 331 (Howard L. MacLeod et al. eds., 2d ed. 2009).

97. *See* IOM, PRIVACY REPORT, *supra* note 3, at 82.

98. *Id.*

99. *See* FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS 35–38 (2010) (warning that the distinction between personally identifiable information and non-identifiable information is increasingly irrelevant in light of the potential for data to be re-identified); Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1706 (2010) (discussing the risks to individual privacy if de-identified data were to be re-identified); Mark A. Rothstein, *Is Deidentification Sufficient to Protect Health Privacy in Research?*, 10 AM. J. BIOETHICS 3, 5 (2010) ("Despite using various measures to deidentify health records, it is possible to reidentify them in a surprisingly large number of cases . . . ."). *But see* Deven McGraw, *Data Identifiability and Privacy*, 10 AM. J. BIOETHICS 30, 30 (2010) ("Using information in less identifiable form greatly reduces risks to privacy . . . ."); Daniel A. Moros and Rosamond Rhodes, *Privacy Overkill*, 10 AM. J. BIOETHICS 12, 13 (2010) ("There is no evidence to suggest, and no obvious reason to suppose . . . that the current protections of deidentified research information are inadequate.").

100. Evans, *supra* note 25, at 624.

101. *See supra* notes 3, 5.

102. *See* Carl H. Coleman, *Rationalizing Risk Assessment in Human Subject Research*, 46 ARIZ. L. REV. 1, 13–17 (2004) (discussing procedural informality of the Common Rule); *see also* Evans, *supra* note 96, at 332 (discussing procedural informality of the waiver provisions of the Common Rule and HIPAA Privacy Rule); Evans, *supra* note 89, at 5 (same); Evans, *supra* note 25, at 624–25 (same).

103. *See* Barbara J. Evans, *Authority of the Food and Drug Administration to Require Data Access and Control Use Rights in the Sentinel Data Network*, 65 FOOD & DRUG L.J. 67, 102–06 (2010) (discussing the role of IRBs in approving access to data under waiver provisions of the Common Rule and HIPAA Privacy Rule).

cial uses without having the government act as an intermediary.[104] Modern examples include private development corporations, which city and state governments sometimes empower to assemble parcels of land for planned redevelopment projects.[105]

The HIPAA and Common Rule waiver provisions are criticized on various grounds,[106] but the fact remains that there are strong justifications for granting private actors at least some power to approve nonconsensual data access. Private delegations of eminent domain power are justified under three circumstances, all of which are present in the area of informational research. First, they make sense when there are holdouts or other strategic barriers to consensual transactions[107] — in other words, when obtaining consent is "impracticable,"[108] which happens to be one of the conditions[109] for granting a waiver of patient consent under the HIPAA Privacy Rule and the Common Rule. Second, private takings are appropriate in situations where justice and efficiency are better served by transferring the taken property to a subsequent private owner rather than to the government.[110] This is the case, for example, if a private-sector research institution has a greater capability for unlocking the scientific and public health potential of the data than government agencies possess.[111] Third, private takings make sense when a repeated pattern of

---

104. *See* Bell, *supra* note 59, at 517 ("[P]rivate takings — that is, takings carried out by nongovernmental actors — have a solid basis in our legal system."); *id.* at 545, 549–50 (providing examples of delegated private takings); Hart, *supra* note 67, at 1116–17 ("[M]ill acts are a well-known illustration of the state's power to direct the transformation of particular pieces of land in America by delegation of its power to private persons.").

105. Bell, *supra* note 59, at 549–50.

106. *See supra* note 102 (citing procedural critiques); *see also* discussion *infra* Part V.B (discussing the absence of a criterion requiring nonconsensual research uses of data to provide public benefits).

107. *See* Bell, *supra* note 59, at 538–40, 543.

108. *See* 45 C.F.R. § 164.512(i)(2)(ii)(B)–(C) (2010) (requiring impracticability both of obtaining consent and of conducting the research without access to the data, before a HIPAA waiver can be granted); *id.* § 46.116(d)(3) (requiring impracticability of conducting the research without a Common Rule waiver).

109. 45 C.F.R. § 164.512(i)(2)(ii) (2010) (stating HIPAA waiver criteria); *id.* § 46.116(d) (stating Common Rule waiver criteria).

110. *See* Bell, *supra* note 59, at 534 (suggesting that the takings power is warranted only if "the government is the preferred owner for reasons of justice or efficiency").

111. *See* U.S. FOOD & DRUG ADMIN., ADVANCING REGULATORY SCIENCE FOR PUBLIC HEALTH, *available at* http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/BudgetReports/UCM207395.pdf ("During the past two decades, extraordinary investments have led to revolutionary advances in the biomedical sciences. However, FDA's scientific expertise and infrastructure have not kept pace with these advances. . . . FDA is unable to fulfill its mission, in part because it lacks modern scientific expertise."); U.S. FOOD & DRUG ADMIN., THE CRITICAL PATH INITIATIVE: REPORT ON KEY ACHIEVEMENTS IN 2009, *available at* http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/UCM221651.pdf at 1 (describing FDA's Critical Path Initiative, which is pursuing a "new paradigm" that is "building partnerships and creating new opportunities for industry and other stakeholders to share expertise and data"); *Pub-*

similar transactions makes government involvement administratively burdensome.[112] This condition arguably is met in the current era of heavy reliance on informational research;[113] the waiver provisions avoid bottlenecks that could arise if the government acted as an intermediary in every nonconsensual research use of data.

Under a property regime, patients' ability to control uses of their data might be very similar to the substantive entitlements they enjoy under the existing federal regulations. There could, of course, be procedural differences, with the property regime imposing higher "due process" costs[114] on nonconsensual uses of data. Yet high due process costs are themselves a factor that tends to justify private delegations of takings power. If patients owned their data, it is quite possible that some scheme of private eminent domain — perhaps resembling the waiver provisions of the federal regulations — would be necessary to address the high due process costs of securing access to data. Based on precedents in the railroad and utility industries,[115] it would not be out of line for the government to delegate eminent domain power to private actors — such as healthcare data environments and research institutions — that are repeatedly involved in transactions to supply or acquire data for use in informational research. It is hard to make a case that data ownership would give patients any more control than they now have.

III. WHY DATA OWNERSHIP CANNOT RESOLVE DATA ACCESS PROBLEMS

Turning to the pro-access proposals, the health information privacy community was puzzled recently by a debate in which several of

---

lic/Private Partnership Program, Biomarker Consortium (BC), U.S. FOOD & DRUG ADMIN., http://www.fda.gov/AboutFDA/PartnershipsCollaborations/PublicPrivatePartnershipProgram/ucm231115.htm (last visited Dec. 21, 2011) (describing a public-private partnership the FDA has formed with twenty-eight private corporations and thirty-four independent research foundations and institutions to accelerate progress in biomarker-based technologies).

112. *See* Bell, *supra* note 59, at 545, 561–62 (providing the example of utility companies that have a need to conduct repeated transactions to acquire rights-of-way).

113. *See* Brian L. Strom, *Preface* to PHARMACOEPIDEMIOLOGY, *supra* note 9, at xvi (noting that epidemiological data are now routinely used in regulatory decision making); IOM, LEARNING HEALTHCARE, *supra* note 42, at 129–30 (discussing the value and challenges of observational methodologies); Evans, *supra* note 42, at 438–39 (attributing the growth of information-based research after 1980 to various stimuli, including improvements in database technology); AHRQ, *Fact Sheet*, *supra* note 18 (discussing the rise and future directions of outcomes research).

114. *See* Merrill, *supra* note 61, at 77 (discussing the procedural complexity of eminent domain, which imposes due process costs in the form of difficulties obtaining legislative authority for a taking, drafting and filing the complaint, serving of process, securing a formal appraisal of the asset's value, and potentially litigating trials and appeals).

115. Bell, *supra* note 59, at 545, 561–63; Hart, *supra* note 67, at 1102.

our most admired scholars drew divergent conclusions about the optimal scheme of ownership for patient data. Divergent conclusions are not puzzling in themselves, but they are so when two analyses that embrace similar objectives, similar methodologies, and similar assumptions give rise to the divergence. Both analyses — one by Professors Hall and Schulman,[116] the other by Professor Rodwin[117] — favor the objective of making health data more widely available for use in medical treatment, public health, and research.[118] Both employ resource classification as their methodology — a method in which analysts "classify infrastructure resources as public goods, network goods, natural monopoly, or some combination thereof"[119] to explain why "markets may fail to efficiently supply such goods, and then proceed to analyze the form of institutional intervention by the government to correct the failure."[120] Both analyses state many of the same assumptions: that the use of health data is nonrivalrous,[121] health data resources generate public goods,[122] interoperable data systems exhibit network effects,[123] data holders such as insurers and healthcare providers enjoy rights somewhat equivalent to ownership of patient data amid the present legal ambiguities,[124] and control of data resources is

116. Professors Hall and Schulman argue in favor of patient ownership of health data. *See* Hall & Schulman, *supra* note 27; *see also* Hall, *supra* note 23 (making similar arguments in a longer legal analysis).

117. Rodwin, *supra* note 23; Rodwin, *supra* note 28 (arguing for public ownership of de-identified patient data).

118. *See* Hall, *supra* note 23, at 635–36 (discussing the major challenge of creating "an interconnected, automated, networked world where information follows the patient, information-based tools aid in decision making, and population health data can be mined to improve the quality and outcome of care for all"); Rodwin, *supra* note 23, at 586–87 (summarizing the advantages of tapping data from patient records in order to "improve medical knowledge, patient safety and public health").

119. Frischmann, *supra* note 75, at 939–40.

120. *Id.* at 929, 939–41.

121. Hall, *supra* note 23, at 661 ("Information by its nature is nonrivalrous, meaning it can be used by many people at once without depletion."); *see* Rodwin, *supra* note 23, at 598 (noting that with public goods, "an individual's use does not diminish use by another person").

122. *See* Rodwin, *supra* note 23, at 597–98, 618; *cf.* Hall, *supra* note 23, at 643 ("[H]oarding medical information destroys the commons that might otherwise support valuable public goods.").

123. Rodwin, *supra* note 23, at 597–98 ("Patient data display network effects."); *see also* Hall, *supra* note 23, at 638 (claiming that network effects emerge by connecting medical records).

124. *See* Hall, *supra* note 23, at 646 (noting that information held by such entities is "out of circulation even though it is not, strictly speaking, owned"); Rodwin, *supra* note 23, at 588 (noting that data holders "treat patient data as if it were their private property"); *id.* at 593 (asserting that "[i]f legislation does not resolve the ownership of data, courts are likely to grant property interests to those who possess [patient] data and preserve the status quo").

highly fractured at the level of these data holders, leading to a tragedy of the anticommons.[125]

Despite their similar approaches, the authors recommend starkly different policy interventions. Rodwin calls for public ownership of patients' anonymized data.[126] He argues "that treating patient data as private property precludes forming comprehensive databases required for many of its most important public health and safety uses" and proposes "that federal law require providers, medical facilities and insurers to report key patient data in anonymized and de-identified form to public authorities, which will create aggregate databases to promote public health, patient safety, and research."[127] Rodwin's proposal is, in effect, a scheme of nonconsensual access to patients' anonymized data, although he does not label it as such. Data holders, such as healthcare providers and insurers, would be required by statute to report data to public authorities.[128] Implicit in this scheme is that patients have no say in the matter. This nonconsensual ordering of data access is a critical feature of Rodwin's proposal.[129]

Hall and Schulman, on the other hand, call for consensual ordering of data access. They suggest that it would stimulate market development of interconnected electronic medical records ("I-EMRs")[130] if patients had a right to enter commercial transactions to license access to their medical information that is in the custody of insurers, healthcare providers, and other data holders.[131] In his longer analysis, Hall argues that the U.S. healthcare system's fragmentation is "chronic and deeply entrenched"[132] such that patients' medical information is widely scattered among data holders who may lack incentives to develop I-EMRs. Patients have rights of access to their own data,[133] but in Hall's view, lack clear entitlements to transfer these rights on commercial terms to "compilers" that could assemble the patient's scattered data into I-EMRs. "If patients were given ownership of their complete medical treatment and health histories, they could license to compilers their rights to that information in a propertized form that could be more fully developed and commercialized."[134]

---

125. *See* Hall, *supra* note 23, at 647 (noting that the data holders' "[m]ultiple ownership of different pieces of a patient's medical history . . . makes it difficult for anyone to assemble a complete record"); Rodwin, *supra* note 23, at 606 (discussing fracturing of control over patient data at the level of physicians, hospitals, and insurers, and noting a second level of fracturing of control at the level of individual patients).

126. Rodwin, *supra* note 28, at 86.

127. Rodwin, *supra* note 23, at 589.

128. *Id.*

129. *See* discussion *infra* Part III.C.

130. *See* Hall, *supra* note 23, at 636 (defining I-EMRs).

131. *Id.* at 638; *see also* Hall & Schulman, *supra* note 27, at 1283–84.

132. Hall, *supra* note 23, at 640.

133. *Id.* at 649–50.

134. Hall, *supra* note 23, at 651.

Hall's analysis is sometimes characterized as a call for private ownership of data.[135] It should not, however, be confused with the demands for property rights sometimes voiced by privacy advocates with the aim of blocking data access.[136] Hall offers a nuanced analysis, recognizing that the precise scope of the patient's entitlement would need to be carefully defined and acknowledging a risk that giving patients additional legal rights could add new strategic barriers in a market that already, in Hall's view, exhibits an anticommons problem at the level of data holders.[137] Hall and Schulman present their proposal as "one potential solution."[138] Under their proposal, the patient would be able to grant a license to a trusted intermediary, which in turn would be able to (1) compel the various data holders to make the patient's medical information available for compilation into an I-EMR (subject, of course, to reimbursing the data holder's costs of complying with such requests)[139] and (2) act as the patient's agent for purposes of arranging commercial transactions with third parties that desire to use the patient's I-EMR.[140] The patient would control the terms under which the trusted intermediary could license the patient's I-EMR to prospective data users, and patients would have a "nonwaivable right to revoke any permission they give for access or use."[141] This scheme of patient-controlled I-EMRs would differ from familiar "ownership of houses and cars."[142]

Both of these analyses are insightful and have advanced the scholarly debate about data ownership, access, and privacy. The comments offered here aim to build on, rather than quibble with, these two proposals. This section examines why neither of the proposals just discussed would fix the data access problem. To explain why these proposals fail, it is necessary to explore some of the implicit assumptions on which they rest. This section focuses on two basic questions that can generate divergent assumptions. The first question is, "What types of health data constitute useful data resources — in

---

135. *See, e.g.*, Rodwin, *supra* note 23, at 608 ("Professor Mark Hall argues that private ownership can overcome anticommons problems that block the adoption of integrated EMRs and networks."); *Who Should Own Electronic Medical Records?*, ALLBUSINESS (2009), http://www.allbusiness.com/print/13276486-1-22eeq.html (labeling the Hall and Schulman analysis as "the case for private ownership" and presenting it as "the opposing view" to Rodwin's "case for public ownership of data").

136. *See* discussion *supra* Part II.

137. *See* Hall, *supra* note 23, at 646–47.

138. Hall & Schulman, *supra* note 27, at 1284.

139. Hall, *supra* note 23, at 650 (noting that the patient's access to patient records held by HIPAA-covered entities is subject to payment of fees to cover the costs of preparing and copying the records, and offering that a "potential solution for the fee problem is insurance reimbursement").

140. *Id.* at 660–61.

141. *Id.* at 661.

142. Hall & Schulman, *supra* note 27, at 1282.

other words, data in a form that can support useful applications in clinical care, public health, and research?" Part III.A defines four categories of health data resources that differ markedly in their utility for these various applications. The second question is, "How are useful data resources made?" Making some types of data resource, as discussed in Part III.B, requires access to identifying information about patients. Part III.C explains that making one particularly useful type of data resource requires a scheme of nonconsensual access to patients' data. Making data resources also requires significant inputs of human and infrastructure services, as described in Part III.D. Simply owning data will not ensure an adequate supply of data resources without access to the necessary services. Proposals that fail to address these realities cannot resolve the data access problem.

A. Identifying the Valuable Data Resources

In ordinary usage, terms like "medical information" and "health data" can refer to several distinct types of information resource:

1. Records of a Patient's Various Encounters with the Healthcare System

This discussion will use the terms "encounter-level patient data" or "raw patient health data" to refer to records of a patient's various encounters with the healthcare system, such as paper charts or electronic files stored by the healthcare providers, payers, clinical laboratories, pharmacies and other sellers of medical products with which the patient has done business in the course of receiving healthcare. Hall's "medical information"[143] and Rodwin's "patient data"[144] often refer to encounter-level patient data. Hall's electronic health records ("EMR") are electronic records of a patient's encounters with a single healthcare site, such as a hospital or a physician's office.[145]

2. The Patient's Longitudinal Health Record ("LHR")

An LHR compiles a patient's encounter-level data from disparate sources to form an extended chronological record that tracks the patient's illnesses, treatments, and outcomes over multiple encounters with the healthcare system.[146] Hall and Schulman refer to these as

143. Hall, *supra* note 23, at 646.
144. Rodwin, *supra* note 23, at 586.
145. Hall, *supra* note 23, at 643–44.
146. *See* Deborah Shatin, Nigel S.B. Rawson & Andy Stergachis, *UnitedHealth Group*, *in* PHARMACOEPIDEMIOLOGY, *supra* note 9, at 271, 273 (discussing "longitudinal histories"

"longitudinal patient records"[147] or "a consolidated medical record for each patient."[148] Hall's I-EMRs would produce LHRs by "facilitat[ing] the compilation of a patient's entire medical treatment and health history from among multiple independent records holders."[149]

3. Longitudinal Population Health Data ("LPHD")

LPHD gathers LHRs from many patients to create a dataset that reflects the long-term healthcare experiences of a large number of people.[150] Hall's trusted intermediaries would be able to enter transactions that gather patient-specific LHRs together to form LPHD. Rodwin's "national patient database" is intended to generate LPHD.

4. Unbiased LPHD

This is a subcategory of LPHD that, in addition to the characteristics just described, has a valuable attribute: the LPHD provides a representative sample[151] of a larger population about which researchers or public health authorities (together, "investigators") are trying to draw conclusions. The individual LHRs included in the unbiased LPHD constitute a representative sample of a larger population of interest. Studying unbiased LPHD will let investigators draw scientifically valid conclusions that are generalizable to that larger population.[152] LPHD obviously is unbiased if it includes LHRs for *all* members of the larger population in question, for example, if it includes data for all Americans or data for everyone in the world. Rodwin's concept of "national, longitudinal patient data"[153] is a form of

---

in which "[i]nformation on diagnosis, treatments, and the occurrence of adverse clinical events, as coded on [insurance] claims, can be tracked across time").

147. Hall & Schulman, *supra* note 27, at 1284.

148. Hall, *supra* note 23, at 635.

149. *Id.* at 651.

150. Evans, *supra* note 25, at 592.

151. *See generally* David M. Eddy, *Should We Change the Rules for Evaluating Medical Technologies?*, *in* IOM, MODERN METHODS, *supra* note 24, at 117, 124–25 (discussing various types of bias that can occur in scientific studies and their impact on generalizability of results).

152. *See* FURBERG & FURBERG, *supra* note 9, at 35–36 (discussing problems that can affect data quality, including biases that can undermine the generalizability of results); *infra* note 173 and accompanying text for discussion of empirical studies demonstrating biases that can result when inclusion of individuals' data into a dataset is predicated on informed consent; *see generally* Brian L. Strom, *Sample Size Considerations for Pharmacoepidemiology Studies*, *in* PHARMACOEPIDEMIOLOGY, *supra* note 9, at 29, 30–35 (discussing the sample sizes required for various types of health informational research); Suzanne L. West, Brian L. Strom & Charles Poole, *Validity of Pharmacoepidemiologic Drug and Diagnosis Data*, *in* PHARMACOEPIDEMIOLOGY, *supra* note 9, at 709 (discussing problems with data quality in health informational research).

153. Rodwin, *supra* note 23, at 587.

unbiased LPHD. Unbiased LPHD also can be created using smaller samples of individuals' LHRs, so long as a random, representative sample is obtained. Hall's statement that "population health data can be mined to improve the quality and outcome of care for all"[154] presumes the use of unbiased LPHD.

Encounter-level patient data and individual LHRs are useful for purposes of treating the individual patient. For example, LHRs can answer questions about a patient's medical history that are relevant to the current treatment encounter, or they can notify physicians about treatments that other care providers have administered to the same patient for the same illness so that duplicative or conflicting treatments can be avoided. Encounter-level patient data and individual LHRs, which include data for just one person, have limited or no direct use as resources for public health studies and research, since it is hard to draw general conclusions from the experiences of one person.[155] However, encounter-level data and LHRs are raw material from which useful data resources for research and public health can be derived. The useful resource for public health and research activities is LPHD.[156] For the vast majority of research and public health studies, there is a further requirement to use unbiased LPHD that can support valid, generalizable scientific conclusions.[157] There are some research and public health applications that can work with biased LPHD. For example, biased LPHD may be useful in preliminary studies to generate hypotheses for later, more rigorous study. In general, however, research and public health studies that use LPHD need *unbiased* LPHD.

---

154. Hall, *supra* note 23, at 635–36.

155. *See* U.S. PREVENTIVE SERVS. TASK FORCE, GUIDE TO CLINICAL PREVENTIVE SERVICES 862 (2d ed. 1996), *available at* http://odphp.osophs.dhhs.gov/pubs/guidecps/PDF/APPA.PDF (comparing the quality of evidence produced by various methodologies and according the lowest ranking to opinions based on reports of outcomes in specific patient cases).

156. *See* PHARMACOEPIDEMIOLOGY, *supra* note 9, at Part III (containing a series of articles describing the types of data that are useful in various types of pharmacoepidemiological studies); U.S. FOOD & DRUG ADMIN., SENTINEL NETWORK PUBLIC MEETING 51–56 (Mar. 7, 2007) (statement of Dr. Marc Overhage) [hereinafter FDA, MARCH 7 PROCEEDINGS], *available at* http://www.fda.gov/ohrms/dockets/dockets/07n0016/07n-0016-tr00001.pdf (discussing the importance and difficulty of linking data longitudinally); U.S. FOOD & DRUG ADMIN., SENTINEL NETWORK PUBLIC MEETING 73–74 (Mar. 8, 2007) (statement of Dr. Clement McDonald) [hereinafter FDA, MARCH 8 PROCEEDINGS], *available at* http://www.fda.gov/ohrms/dockets/dockets/07n0016/07n-0016-tr00002.pdf (discussing the importance of longitudinal population health data in research and noting the difficulties of linking data from disparate data sources).

157. *See* discussion *infra* Part III.C.

B. The Problem of Linking Data Across Healthcare Data Environments

Rodwin calls for encounter-level patient data to be "anonymized or de-identified"[158] at the source by each data holder, which then would report the anonymized data to a centralized, national database that would somehow "create aggregate databases to promote public health, patient safety, and research."[159] This proposal runs into a serious technical problem: it is impossible — not merely costly or difficult, but impossible — to make longitudinal health records out of encounter-level patient data that have been anonymized.[160] Linking data longitudinally to create a patient's LHR requires at least some identifying information to establish that raw data received from various data holders relate to the same patient.[161] If the goal is to make an anonymized LHR, the order of operations matters: first, identifiable encounter-level patient data are linked together to make an identifiable LHR; then the identifiable LHR is anonymized.[162] The linkage must precede the anonymization.

Rodwin's proposal would let each data-holding facility report its data in coded[163] form — that is, with an individual tracking number that allows data from the patient's subsequent encounters with that facility to be linked to data already reported.[164] Unfortunately, these facility-level tracking numbers would not let a patient's data from one facility be linked with data from *other* facilities where the patient has received care.[165] Each facility maintains its own coding system,[166] and a patient would be assigned different tracking numbers by the various facilities with which she interacts. Sharing of code keys, which relate the tracking numbers to specific individuals, amounts to sharing of

158. Rodwin, *supra* note 23, at 589.

159. *Id.*

160. *See* FDA, MARCH 7 PROCEEDINGS, *supra* note 156, at 51–56 (statement of Dr. Overhage) (discussing the process of linking data longitudinally and noting the necessity for some sharing of identifiable information); FDA, MARCH 8 PROCEEDINGS, *supra* note 156, at 73–74 (statement of Dr. MacDonald) (discussing the need for sharing of identifiable information to accomplish linkage); Evans, *supra* note 25, at 594–96, 606.

161. Evans, *supra* note 25, at 594–96, 606.

162. *See* Evans, *supra* note 103, at 77 fig.2 (explaining that longitudinal linkage requires the use of at least some identifying information, such that de-identification must follow, rather than precede, linkage).

163. *See generally* Evans, *supra* note 25, at 619–31 (discussing coding of data and its significance under the HIPAA Privacy Rule, the Common Rule, and the FDA human subject protection regulations).

164. Rodwin, *supra* note 23, at 615.

165. *See* Evans, *supra* note 103, at 76–77 (discussing the difficulties with linking data across multiple healthcare data environments).

166. FDA, MARCH 7 PROCEEDINGS, *supra* note 156, at 52–53 (statement of Dr. Marc Overhage) ("Every institution typically has some kind of unique identifier for the individual patient . . . [that] isn't linked to anything else.").

identifiable data under the Privacy Rule and Common Rule.[167] Rodwin's proposal, which only allows sharing of de-identified patient data, thus would not allow sharing of the code keys.

Suppose Mary Smith visits Dr. Brown for arthritis pain and is prescribed rofecoxib (Vioxx). Two months later, she is treated at the emergency department of Central Hospital for a stroke. Six months later, she visits Dr. Brown again for a skinned knee. Under Rodwin's proposal, the national database would contain the following information: Dr. Brown treated anonymous patient 13275 for arthritis and prescribed Vioxx. This same patient (identified by Dr. Brown's tracking number 13275) was later treated for a skinned knee. Central Hospital treated anonymous patient 999345 for a stroke. There is no way to link Mary's data from Dr. Brown and Central Hospital into a complete LHR unless they divulge that tracking numbers 13275 and 999345 both refer to Mary Smith. If the data in the national database are "mined" for information about Vioxx safety, investigators will see a possible association between taking Vioxx and skinning one's knee, but they will not be able to detect the possible association between taking Vioxx and having a stroke.

Under Rodwin's proposal, the national database would not be able to compile LHRs for each patient, since it would lack the identifying information needed to link the patient's encounters across multiple facilities and data holders. Unable to create LHRs, the national database could not produce the "[n]ational, longitudinal patient data"[168] (LPHD) that are needed for public health activities and research. The proposed national database would merely contain a second, anonymized copy of the same fragmented, unlinked, disorganized data that already exist. Unless encounter-level patient data are shared with the government in identifiable form — a policy that is far more problematic[169] than the one Rodwin has proposed — it is difficult to see how a national database would add any value.

---

167. *See* 45 C.F.R. § 164.514(b)(2)(i)(R), 164.514(c)(2) (2010) (stating that "de-identification" of data under the HIPAA Privacy Rule requires removal of codes, but making an exception for codes that comply with the standard set in section 164.514(c), which forbids disclosure of the "mechanism for re-identification" (i.e., the code key)); *see also OHRP Guidance*, *supra* note 10 (noting that if code keys are shared with investigators, the study will be considered human-subject research that requires consent under the Common Rule).

168. *See* Rodwin, *supra* note 23, at 587.

169. *See* IOM, PRIVACY REPORT, *supra* note 3, at 82 (reporting results of multiple surveys that found "[p]atients were much more comfortable with the use of anonymized data (e.g., where obvious identifiers have been removed) than fully identifiable data for research").

C. Consent Bias and the Need for Nonconsensual Access to Patients' Health Data

The proposal by Hall and Schulman solves the data-linkage problem by relying on consensual ordering.[170] The patient could authorize a trusted intermediary to obtain identifiable encounter-level data, which then could be linked to create the patient's LHR. The resulting LHR would be useful for purposes of the patient's own care.[171] It is not clear, however, that this proposal could generate data resources for research and public health activities. The problem relates to the consensual ordering inherent in Hall and Schulman's scheme of patient-controlled health records. The Hall and Schulman proposal allows the trusted intermediary to use a patient's LHRs to form LPHD and to license the LPHD to third-party users — *but only on terms controlled by the patient*.[172]

Multiple empirical studies have documented that people who are willing to consent to letting their data be used in research differ *medically* from the population at large.[173] The underlying reasons are not well understood, but the impact is clear: conditioning the creation of LPHD on patient consent produces datasets that may be unreflective of the general population, thus biasing study results. Similar problems also exist outside the biomedical context. Burstein notes that information security researchers, when studying threats to our nation's critical information infrastructures, need access to realistic data about people's Internet usage patterns and electronic communications.[174] Internet service providers that possess this information can share it with researchers, but only if the affected Internet users consent.[175] People willing to consent to research with their private information — potentially including the content of their e-mails — may not provide a

---

170. *See* Hall & Schulman, *supra* note 27, at 1284 (calling for "[p]atient-controlled health records").

171. *See* discussion *supra* Part III.A.

172. Hall, *supra* note 23, at 660–61.

173. *See generally* Brian Buckley et al., *Selection Bias Resulting from the Requirement for Prior Consent in Observational Research: A Community Cohort of People with Ischaemic Heart Disease*, 93 HEART 1116 (2007); Casarett et al., *supra* note 9, at 593–94; IOM, PRIVACY REPORT, *supra* note 3, at 209–14 (surveying studies of consent and selection bias); Khaled El Emam et al., *A Globally Optimal k-Anonymity Method for the Deidentification of Health Data*, 16 J. AM. MED. INFO. ASS'N. 670, 670 (2009); Steven J. Jacobsen et al., *Potential Effect of Authorization Bias on Medical Record Research*, 74 MAYO CLINIC PROC. 330 (1999); Jack V. Tu et al., *Impracticability of Informed Consent in the Registry of the Canadian Stroke Network*, 350 NEW ENG. J. MED. 1414 (2004); Steven H. Woolf et al., *Selection Bias from Requiring Patients to Give Consent to Examine Data for Health Services Research*, 9 ARCHIVES FAM. MED. 1111 (2000).

174. Aaron J. Burstein, *Amending the ECPA to Enable a Culture of Cybersecurity Research*, 22 HARV. J.L. & TECH 167, 170–71, 184–94 (2008).

175. *Id.* at 185–86 (citing the Omnibus Crime Control and Safe Streets Act of 1968 and the Electronic Communications Privacy Act of 1986).

representative sample of all Internet users and likely would not include the cyberterrorists that researchers were hoping to study.

Because of consent bias, Hall and Schulman's patient-controlled I-EMRs can generate individual patients' LHRs but cannot produce the high-quality, unbiased LPHD that researchers and public health officials need in order to draw scientifically valid conclusions. The Hall and Schulman proposal envisions that licensing fees paid by third-party data users would help finance the informational infrastructure for compiling patients' LHRs.[176] Given the low quality of LPHD a patient-controlled system can generate, demand from research and public health users may be limited, and licensing fees may not be a reliable source of funding for the system.[177]

Consent bias is a potential problem under any scheme of consensual ordering, and this is true whether it is an opt-in or opt-out consent scheme. An opt-in approach allows data to be used only if the patient affirmatively consents.[178] Common Rule consents and HIPAA privacy authorizations exemplify an opt-in approach.[179] An opt-out approach presumes patients' data can be used, unless the patients take active steps to exclude their data from use.[180] Both schemes let patients exercise control over uses of their data, although the level of effort involved in keeping one's data from being used differs in the two schemes. Whenever patients can control uses of their data, there is a risk that those who exercise control may differ from those who do not, causing the resulting data set to be unrepresentative of the population as a whole. A scheme of nonconsensual ordering avoids this problem because patients cannot self-select for inclusion or exclusion from the data set.

Rodwin's analysis excels in its exposition of supply-side factors that call for nonconsensual ordering of access to data for research and

---

176. Hall & Schulman, *supra* note 27, at 1283 (suggesting that patients could authorize intermediaries to sell their data for use in marketing and research to permitted users, with proceeds helping to "recoup the considerable expenses of compilation" of the data and possibly providing some flow of funds back to the patient); *see* Hall, *supra* note 23, at 646 ("[P]ropertizing medical information could stimulate increased flow of medical information into more useful forms by giving stakeholders rights that they can license or sell.").

177. Hall and Schulman never expressly claimed that their proposed scheme would produce data suited to research and public health uses; they may have envisioned I-EMRs primarily as a tool to improve clinical care. *See* Hall, *supra* note 23, at 650 (suggesting that health insurers might be a source to help cover the costs of generating patient's interconnected EMRs — a notion that seems to presume I-EMRs would be used in clinical care rather than in research and public health uses).

178. Mark A. Rothstein, *Health Privacy in the Electronic Age*, 28 J. LEGAL MED. 487, 490–91 (2007).

179. *See* 45 C.F.R. § 46.116 (2010) (describing Common Rule consent requirements); *id.* § 164.508 (describing HIPAA authorization requirements).

180. *See, e.g.*, Michael Birnhack & Niva Elkin-Koren, *Does Law Matter Online? Empirical Evidence on Privacy Law Compliance*, 17 MICH. TELECOMM. & TECH. L. REV. 337, 339 (2011), *available at* http://www.mttlr.org/volseventeen/birnhack_elkin-korens.pdf.

public health applications.[181] Hall and Rodwin both acknowledge that diffusion of control among multiple data holders can give rise to a tragedy of the anticommons.[182] Rodwin explores an additional tragedy of the anticommons that arises when control over data is diffused at the level of individual patients.[183] Likening the problem of assembling "comprehensive patient databases" to the problem of assembling contiguous parcels of land for real estate development, he explores strategic barriers that make consensual access to data unworkable.[184]

Rodwin frames his discussion as a comparison of private and public ownership. This framing obscures an essential feature of his proposal: it is a scheme of nonconsensual access to patients' data, insofar as it requires compulsory reporting of patients' data to the government.[185] The flaw in Rodwin's analysis is that it conflates nonconsensual access and governmental ownership. He states that "treating patient data as private property precludes forming comprehensive databases required for many of its most important public health and safety uses."[186] This statement is true only if property rights are modeled as conferring property-rule protection (pure consensual ordering). It discounts the possibility that the needed public access to privately owned data could be obtained nonconsensually through exercises of the police or eminent domain powers.[187] As Bell has remarked in discussing takings that transfer property into public ownership, such actions are "warranted only where two issues are resolved in favor of the government: (1) the government is the preferred owner for reasons of justice or efficiency, and (2) coercion is the preferred transfer mechanism."[188] The need for nonconsensual access does not necessarily imply a need for governmental ownership.

There are various reasons why the government may not be the most efficient data owner. For example, federal agencies such as the HHS, which would own the data under Rodwin's proposal,[189] are regulated by the Privacy Act,[190] which protects the privacy of records held by federal agencies.[191] This is an added layer of regulation on top of the HIPAA Privacy Rule and the Common Rule. This heightened

181. *See* Rodwin, *supra* note 23, at 603–06.

182. Hall, *supra* note 23, at 647–48; Rodwin, *supra* note 23, at 606.

183. Rodwin, *supra* note 23, at 606.

184. *Id.* at 607; *see also* Hall, *supra* note 23, at 647 (invoking the land-assembly analogy to describe strategic barriers in getting multiple data holders to cooperate to assemble a patient's complete longitudinal health record).

185. *See* Rodwin, *supra* note 23, at 589.

186. *Id.*

187. *See* discussion *supra* Part II.A.

188. Bell, *supra* note 59, at 534.

189. *See* Rodwin, *supra* note 28, at 86; Rodwin, *supra* note 23, at 615.

190. 5 U.S.C.A. § 552(a) (West 2006 & Supp. 2011); *see also* Rosati, *supra* note 89, at 5.

191. IOM, PRIVACY REPORT, *supra* note 3, at 89.

regulatory burden, along with other potential disadvantages of public ownership, would need to be carefully weighed, even if one is prepared to accept that HHS has the resources to construct and operate a mega-database containing duplicate copies of all of the health data in the United States. It is thus unclear whether public data ownership is a good idea. On the other hand, there is a strong case for nonconsensual access to data for at least some research and public health applications — specifically, those that require unbiased LPHD. Because of the patient-level anticommons problem Rodwin explored[192] and concerns about consent bias, consensual approaches cannot reliably produce unbiased LPHD.

D. The Role of Infrastructure and Demand-Side Factors

Statements such as "[w]hoever owns patient data will determine whether its benefits can be tapped"[193] overstate the importance of controlling one raw material input to a complex, multistage production process. This statement is true in the same way that the statement "whoever owns iron ore will determine the fate of the shipbuilding industry" is true. Certainly, iron ore is a critical input to building a ship, but it is just one of many factors that influence the development of facilities that turn iron ore into a valuable asset — steel — and the steel into ships. In the same way, raw health data are just one of many inputs for creating useful data resources. This Part explains the importance of other critical inputs — specifically, human and infrastructure services.

There are multiple system architectures that can convert encounter-level patient data into valuable data resources for research and public health.[194] Hall's proposal does not embrace any particular system architecture for implementing I-EMRs. He merely states that "[t]he primary barriers are not technological"[195] and turns to analysis of the perceived legal barriers. Rodwin's analysis implicitly assumes that a centralized database is necessary in order to assemble encounter-level patient data into LHRs and LPHD.[196] His preference for public ownership may have been influenced by the assumption — which

192. *See supra* notes 183, 185, and accompanying text.

193. Rodwin, *supra* note 23, at 587.

194. *See* HEALTHCARE INFO. AND MGMT. SYS. SOC'Y, A HIMSS GUIDE TO PARTICIPATING IN A HEALTH INFORMATION EXCHANGE 15–20 (2009), *available at* http://www.himss.org/content/files/HIE/HIE_GuideWhitePaper.pdf (discussing an array of possible architectures, including centralized, decentralized (federated), and hybrid models); Carol C. Diamond, Farzad Mostashari & Clay Shirky, *Collecting and Sharing Data For Population Health: A New Paradigm*, 28 HEALTH AFFAIRS 454, 456 (2009).

195. Hall, *supra* note 23, at 636.

196. Rodwin, *supra* note 23, at 595 (taking the position that "tapping the real potential for patient data for secondary uses requires that it be aggregated into a national database").

is erroneous — that public access to, and use of, data requires actual possession of the data in a centralized database.

It is true that in the past informational research typically was performed by gathering data into one large, central database where the data analysis was performed.[197] The modern trend is to use distributed data networks instead.[198] Centralized databases worked satisfactorily in the days — not so long ago — when a "large scale" observational study might have involved mere tens to hundreds of thousands of records. Today, however, large-scale studies may use records of tens to hundreds of millions of persons.[199] For example, the Food and Drug Administration Amendments Act of 2007 ("FDAAA")[200] calls for pharmacoepidemiological[201] studies of postmarket drug safety that will employ health data for one hundred million persons.[202] FDA is meeting this mandate by developing the Sentinel system,[203] and its pilot Mini-Sentinel system[204] already includes data for sixty million persons.[205] Multimillion-person pharmacoepidemiological networks also are being developed in Canada,[206] the European Union,[207] and

197. Diamond et al., *supra* note 194, at 456.

198. Richard Platt et al., *The New Sentinel Network — Improving the Evidence of Medical-Product Safety*, 361 NEW ENG. J. MED. 645, 645–47 (2009); *see also* Diamond et al., *supra* note 194, at 460.

199. *See* Evans, *supra* note 103, at 73–74 (describing several multimillion-person pharmacoepidemiological data networks now under development).

200. Pub. L. No. 110-85, 121 Stat. 823 (2007) (codified as amended in scattered sections of 21 U.S.C.).

201. *See* Brian L. Strom, *What is Pharmacoepidemiology?*, *in* PHARMACOEPIDEMIOLOGY, *supra* note 9, at 3, 3 (defining pharmacoepidemiology as "the study of the use of and the effects of drugs in large numbers of people").

202. 21 U.S.C.A. § 355(k)(3)(B)(ii) (West 2006 & Supp. 2011) (setting targets of twenty-five million persons by July 2010 and 100 million by July 2012); *see also id.* § 355(k)(3)(C) (describing the new "postmarket risk identification and analysis system").

203. U.S. FOOD & DRUG ADMIN., THE SENTINEL INITIATIVE (2008), *available at* http://www.fda.gov/downloads/Safety/FDAsSentinelInitiative/UCM124701.pdf (discussing the goals and structure of the Sentinel data network); *see also FDA's Sentinel Initiative*, U.S. FOOD & DRUG ADMIN., http://www.fda.gov/Safety/FDAsSentinelInitiative/default.htm (last modified Oct. 5, 2011) (providing information about the current status of Sentinel System development).

204. Press Release, U.S. Food & Drug Admin., FDA Awards Contract to Harvard Pilgrim to Develop Pilot for Safety Monitoring System (Jan. 8, 2010), http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm196968.htm.

205. Rachel E. Behrman et al., *Developing the Sentinel System — A National Resource for Evidence Development*, 364 NEW ENG. J. MED. 498, 498 (2011).

206. *See In Brief: The Drug Safety and Effectiveness Network (DSEN)*, CANADA INSTS. OF HEALTH RES., http://www.cihr-irsc.gc.ca/e/39389.html (last modified June 29, 2011) (describing Canada's DSEN network); *Medicines that Work for Canadians: Business Plan for a Drug Effectiveness and Safety Network*, HEALTH CANADA (2007), http://www.hc-sc.gc.ca/hcs-sss/alt_formats/hpb-dgps/pdf/pubs/pharma/2007-med-work_eff/2007-med-work-eff-eng-final.pdf (same).

207. *See* Press Release, European Medicines Agency, EMEA-Coordinated PROTECT Project Has Been Accepted for Funding by the Innovative Medicines Initiative Joint Undertaking (Apr. 30, 2009), http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2009/11/news_detail_000096.jsp&jsenabled=true (describing the

Japan.[208] These systems have not required any clarification of data ownership and they did not require creation of centralized databases.[209] They rely on distributed network architectures.[210]

In a distributed network, an individual's health information is not moved to a central database for storage and analysis; rather, it continues to be stored at its original location (for example, in an insurer's or healthcare provider's database).[211] The participating data holders are linked together virtually.[212] Under one design, parties wishing to use the data send queries to the data holders.[213] Suppose, for example, that an investigator wishes to study whether taking statins may be associated with rhabdomyolysis, a muscle-wasting condition. The investigator would send queries to the various data holders (for example, "Please locate records for any person in your data system who (1) has ever taken statins, or (2) has ever suffered from rhabdomyolysis."). Records for such patients could be conveyed in identifiable form to a network coordinating center (a trusted intermediary) that would perform longitudinal linkage of data received from the various data holders. This linkage would make it possible to identify patients who both took statins *and* suffered rhabdomyolysis, even if the records of these two occurrences are scattered across multiple data holders. The trusted intermediary would use the linked data to compile lists of patients who took statins with and without subsequently developing rhabdomyolysis. These lists then could be de-identified and conveyed to the investigator for use in the study.

PROTECT network); *Implementation of the Action Plan to Further Progress the European Risk Management Strategy: Rolling Two-Year Work Programme (2008–2009)*, EUROPEAN MEDICINES AGENCY (Dec. 24, 2007), http://www.emea.europa.eu/pdfs/human/phv/28008907en.pdf (describing the ENCePP data network); EUROPEAN NETWORK OF CENTRES FOR PHARMACOEPIDEMIOLOGY AND PHARMACOVIGILANCE ("ENCePP"), http://encepp.eu (last modified Sept. 30, 2011); EU-ADR, http://www.alert-project.org/ (last visited Oct. 10, 2011) (describing the European Union adverse drug reactions data network).

208. Kaoru Misawa, Dir., Office of Safety, Pharm. & Med. Devices Agency ("PMDA"), Address at the 9th Kitasato University-Harvard School of Public Health Symposium: Sentinel Initiative in Japan: Utilization of Electronic Health Information in Pharmacovigilance 7–14 (Sept. 11–12, 2009), http://www.pharm.kitasato-u.ac.jp/biostatis/khsympo200909/doc/misawa.pdf.

209. *See generally* Behrman et al., *supra* note 205 (discussing development of the Sentinel System). The Sentinel System has been implemented within the framework of existing state law without changes to data ownership arrangements. *See* Richard Platt et al., *supra* note 198, at 645–46 (describing the Sentinel System's distributed architecture); *supra* notes 206, 207 (briefly describing the architectures of related systems in Canada and Europe).

210. *See supra* note 209.

211. Evans, *supra* note 103, at 76.

212. *Id.* at 75–78 (discussing distributed architectures).

213. *Id.* at 77 fig.2 (showing a distributed network query structure that provides for longitudinal linkage of data across participating data environments via a trusted intermediary).

Distributed architecture offers a number of advantages over central databases.[214] It avoids the need to invest in duplicative storage capacity because data reside with the original data holders and are not redundantly stored at a central location. It also offers advantages in privacy and data security because the data continue to reside behind the privacy firewalls of the original data holders, with movements of data minimized to what is necessary to respond to specific queries (as opposed to moving all the data to a central repository in anticipation of unspecified future uses).[215] Perhaps the most important advantage, in terms of data quality, is that distributed networks allow the encounter-level data to be interpreted and processed by the data holders' own personnel, who regularly work with the data and are familiar with its quirks.[216] Data holders do not all use standardized record formats.[217] Different healthcare providers and insurers describe the same medical condition in different ways, just as law professors use different terminology to refer to similar concepts (for example, LHR, I-EMR, complete patient record, and longitudinal patient data). Answering a simple question, such as whether a patient actually had rhabdomyolysis, requires familiarity with how the particular data system records data. The President's Council of Advisors on Science and Technology ("P-CAST") is pessimistic that a standard record format will ever emerge: "[A]ny attempt to create a national health IT ecosystem based on standardized record formats is doomed to failure . . . . With so many vested interests behind each historical system of recording health data, achieving a natural consolidation around one record format . . . would be difficult, if not impossible."[218] The notion that a national database operator could make sense of raw, encounter-level patient data reported in disparate formats is fanciful.

In a distributed data network, data holders are not just suppliers of data; they also act as service providers.[219] These services may include, for example, searching the data holders' records to locate data relevant to the particular query, retrieving data, converting data to a

---

214. *See* Diamond et al., *supra* note 194, at 459; Platt et al., *supra* note 198, at 645 (discussing these advantages).

215. *See supra* note 214; Judith Racoosin et al., Symposium at the 27th International Conference on Pharmacoepidemiology, FDA's Mini-Sentinel Program to Evaluate the Safety of Marketed Medical Products: Progress and Direction 27–28 (Aug. 17, 2011), http://www.mini-sentinel.org/work_products/Publications/Mini-Sentinel_Progress-and-Direction.pdf (listing reasons for preferring a distributed architecture).

216. *See* Platt, *supra* note 198, at 646; Racoosin et al., *supra* note 215, at 28.

217. PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., EXEC. OFFICE OF THE PRESIDENT, REPORT TO THE PRESIDENT: REALIZING THE FULL POTENTIAL OF HEALTH INFORMATION TECHNOLOGY TO IMPROVE HEALTHCARE FOR AMERICANS: THE PATH FORWARD 39 (2010) [hereinafter P-CAST REPORT].

218. *Id.*

219. *See* Evans, *supra* note 103, at 86–90 (discussing the types of infrastructure that FDAAA envisions will be necessary to support operations of FDA's Sentinel System).

common format that will allow data from multiple data holders to be combined, and preparing the search results for transmittal to the user.[220] The term "data provisioning" is sometimes used to refer to these types of services.[221] Encounter-level patient data are transformed into valuable information resources (LHRs and LPHD) through the addition of services.[222]

This fact has important implications. It no longer can be said that "[w]hoever owns patient data will determine whether its benefits can be tapped."[223] Tapping the benefits requires both data and services, and control over data is unavailing without the services. Some argue that health data resources are nonrivalrous.[224] It is probably fair to say that *encounter-level* patient data are nonrivalrous because many people can use raw data without exhausting the supply. However, these data have few uses except in the patient's own care, so it is not clear why large numbers of people would want to use them. LHRs and LPHD, which do have many potential uses, are subject to potential supply constraints: there is a finite supply of the services needed to convert raw data into LHRs and LPHD. Data holders do not have unlimited personnel and data processing resources to respond to queries. Preparing LPHD to respond to one query may diminish the availability of LPHD for another query. The most valuable information resources for clinical, research, and public health applications are LHRs and LPHD, and these can only be supplied by a constrained infra-

220. Houtan Aghili, Senior Technical Staff Member, Presentation to Maryland Task Force: IBM Healthcare & Life Sciences, IBM NHIN-Enabled Health Information Exchange ("NHIE") 3 (July 9, 2007), http://mhcc.maryland.gov/electronichealth/presentations/ibm2_0707.pdf (listing, in a presentation about development of a statewide health information exchange, various "data services" that a health information network needs to enable as "core services," including "secure data delivery;" "data look-up, retrieval, and location registries;" and "data anonymization").

221. *See, e.g.*, *id.* (listing among the core services that a networked health information exchange provides, "[s]upport for secondary use of clinical data including data provisioning"); Paul J. Ambrose, Arun Rai & Arkalgud Ramaprasad, *Internet Usage for Information Provisioning: Theoretical Construct Development and Empirical Validation in the Clinical Decision-Making Context*, 53 IEEE TRANSACTIONS ON ENGINEERING MGMT. 112 (2006), *available at* http://ieeexplore.ieee.org/arnumber=1580898 (providing another example of the term "information provisioning" to refer to making information available for use by decision makers in the healthcare context); *29 Information Provisioning Concepts*, ORACLE STREAMS CONCEPTS AND ADMINISTRATION 11G RELEASE 1 (11.1), http://download.oracle.com/docs/cd/B28359_01/server.111/b28321/strms_ipro.htm #BHCIEBGD (last visited Dec. 21, 2011) ("Information provisioning makes information available when and where it is needed.").

222. *See* Aghili, *supra* note 220, at 3 (providing examples of core services provided by a health information network that is capable of accessing and manipulating raw health data to create useful data resources for secondary purposes such as research and public health applications).

223. Rodwin, *supra* note 23, at 587.

224. *See* Hall, *supra* note 23, at 661 ("Information by its nature is nonrivalrous . . . .").

structure. These resources are only partially nonrivalrous — that is, they are nonrivalrous only within capacity constraints.[225]

The fact that the necessary services are costly and in finite supply has ramifications for system design. A key design decision is whether a system needs to be able to produce LHRs and LPHD ahead of demand, as opposed to satisfying demand after it arises. The answer depends on whether the planned applications — clinical care, research, and public health studies — are latency-sensitive.[226] The concept of latency (colloquially understood as "delay") has been a concern in discussions of Internet policy.[227] Some Internet applications are latency-sensitive — that is, small delays in delivery of information will disrupt their functionality — while others are latency-insensitive. "Consider that it doesn't matter whether an email arrives now or a few milliseconds later. But it certainly matters for applications that want to carry voice or video."[228] Clinical uses of LHRs are potentially latency-sensitive: clinicians treating a patient in the emergency department cannot afford to wait for compilation of the patient's LHR. On the other hand, the use of LHRs in scheduled clinical care may not be latency-sensitive: when a patient makes a doctor's appointment, a request could be made to compile the patient's LHR for delivery on the date of the scheduled appointment. Many research and public health uses of LPHD are latency-insensitive: it does not destroy the validity of a study if it takes a few days or weeks to supply the necessary data resources.

Because of these differences, the optimal infrastructure to supply data resources for one use may not be optimal for supplying other uses. For latency-sensitive applications, data resources need to be compiled ahead of the demand for them. Patient-controlled I-EMRs, such as those proposed by Hall and Schulman, are thus a potentially useful tool for clinical care. Patients can request compilation of their LHRs in advance so that they will be available in emergencies, and then patients can periodically update their LHRs. For latency-insensitive applications, such as most research and public health studies, compilation can be deferred until there is an identified demand. This distinction affects the required system design and can drastically affect system costs when, as here, compiling the information resources requires inputs of scarce, costly services. There would be little advantage — and an enormous cost disadvantage — in developing a centralized, national database containing every person's compiled

225. *See* Frischmann, *supra* note 75, at 951 (defining partially nonrivalrous resources).

226. *See* Tim Wu, *Network Neutrality, Broadband Discrimination*, 2 J. ON TELECOMM. & HIGH TECH. L. 141, 148 (2003) (defining and discussing the impact of latency on Internet applications).

227. *Id.*; *see also* Frischmann, *supra* note 75, at 1008–10.

228. Wu, *supra* note 226, at 148.

LHR. Compiling information resources (LHRs and LPHD) in anticipation of all conceivable research and public health uses may be as ill-advised as it would be to manufacture false teeth for every American in anticipation that they may eventually need them.

A distributed architecture can respond to queries as they occur, and this attribute offers important economic advantages in latency-insensitive research and public health applications. In the future, the development of new infrastructure may reduce the latency itself — in other words, reduce the delays associated with locating relevant patient data and converting them to a consistent format for assembly into LHRs and other useful data resources.[229] At present, these services are labor-intensive. The recent P-CAST report calls for creation of a universal exchange language and infrastructure to facilitate assembly and sharing of patient data across data holders.[230] Data holders would continue to operate a variety of systems, including the old legacy systems in operation today and new recordkeeping systems and formats.[231] The "syntax for such a universal exchange language will be some kind of extensible markup language (an XML variant, for example) capable of exchanging data from an unspecified number of (not necessarily harmonized) semantic realms."[232] Individual data elements — such as a person's X-ray or clinical observations about the patient — would be annotated with metadata tags containing enough identifying information to let the patient's records be located, recording information about the patient's privacy preferences, and explaining the provenance of the data (such as which healthcare providers were involved and what type of test or equipment they used).[233] A national infrastructure would support searches and deliver results appropriately compiled and processed to protect privacy. Locating all of a patient's data, wherever stored, would work similarly to the way an Internet search engine works today.[234] Until such a solution is implemented, LHRs and

229. *See* P-CAST REPORT, *supra* note 217, at 11 (noting that the present "lack of data exchange also means that researchers and public health agencies have limited access to data that could be used to improve health systems and advance biomedical research"); *id.* at 63 (noting that "[t]oday's clinical research studies are not carried out in real time" and may be "[o]ut of date before they are even finished"); *id.* at 54 (calling for creation of a distributed network that "links healthcare providers, patients, labs, researchers, and other stakeholders and enables qualified users to query distributed data stored by partners in the network"); *id.* at 64 (noting that such a system offers the "[p]otential for [r]eal-[t]ime, [r]eal-[w]orld, and [c]omprehensive [d]ata" and discussing various public health and research questions that could be addressed "using large datasets gathered through ongoing medical care, particularly if the data were available in near real time").

230. *Id.* at 4.

231. *Id.* at 41.

232. *Id.*

233. *Id.*

234. *See id.* at 4 (noting that the Office of the National Coordinator for Health Information Technology proposed using the clinical document architecture standard, an estab-

LPHD will continue to require labor-intensive services. The hope, eventually, is to replace some of the human services with more capital-intensive infrastructure services.

Installing new infrastructure requires money. P-CAST acknowledges that federal leadership will be required; "market forces are unlikely to generate appropriate incentives for the necessary coordination to occur spontaneously."[235] This view is far more pessimistic than the view, expressed by Hall and Schulman, that altering patient's entitlements to their health data "will help stimulate market development of interconnected EMRs."[236] The problem with clarifying ownership of health data is that it is a supply-side solution — and this remains true whether ownership is clarified in favor of patients (as in Hall and Schulman's proposal) or the public (as in Rodwin's proposal). In contrast, health information infrastructure — like any infrastructure — exhibits problems both on the supply and demand sides.[237] An example is the P-CAST proposal just described. The proposal could reduce delays in supplying LHRs and LPHD, but the market may not value the incremental speed because many health data applications are not latency-sensitive. Researchers who can afford to wait for a good data set may not be willing to pay more for a good data set delivered sooner. That is a demand-side issue.

There are other demand-side issues. The price users would be willing to pay for health data resources may not reflect the true value of those resources because so many uses of health data (such as research and public health activities) themselves produce public and nonmarket goods.[238] In this situation, data users are unable to appropriate the full value their activities create; thus, they cannot afford to pay a price that reflects the data's true value.[239] Frischmann has noted that market failure for infrastructure is more complex than supply-side analysis suggests.[240] "For both traditional and nontraditional infrastructure resources, analysts emphasize supply-side issues . . . and

---

lished and highly developed technology that is the basis of web search engines, to support indexing and retrieval of metadata-tagged health data across large numbers of geographically diverse locations); *id.* at 42 (indicating that the data-element access services in P-CAST's proposed system "would act much like today's web search engines" but with additional privacy protections).

235. *Id.* at 4.

236. Hall, *supra* note 23, at 638.

237. *See* Frischmann, *supra* note 75, at 930 (arguing that market failures affecting infrastructure industries are complex and include demand-side issues as well as supply-side issues).

238. *Id.* at 966–67 (defining and comparing public and nonmarket goods).

239. *Id.* at 968 ("Infrastructure users that produce public goods and nonmarket goods suffer valuation problems because they generally do not fully measure or appropriate the (potential) benefits of the outputs they produce and consequently do not accurately represent actual social demand for the infrastructure resource.").

240. *Id.* at 930.

assume that the market mechanism will best generate and process demand information."[241] Data propertization proposals assume that if encounter-level patient data were simply assigned to the right owner, the market would be able to figure out the right price to pay for useful data resources such as LHRs and LPHD, and this price would cover the cost of necessary infrastructure and services to create those resources. This is not a safe assumption.

E. Why Data Propertization Proposals Fail

To summarize, encounter-level patient data are an input that can be transformed into high-valued data resources — LHRs and LPHD — for use in clinical care, research, and public health activities. Making these data resources also requires inputs of human and infrastructure services — that is, data provisioning services. In theory, it is possible to produce LHRs for use in clinical care under a patient-controlled system. Such a system would subject all transfers of encounter-level patient data to consensual ordering, which would require permission of the patients whose data are involved. There are major limitations to such a system, however. Because of consent bias, the system cannot supply unbiased LPHD for use in research and public health projects. Research and public health users thus cannot be counted on to cross-subsidize the costs of developing patient-controlled LHRs. Unless the costs of developing patient-controlled LHRs are justified by the value they create in clinical care, a patient-controlled system may not be financially viable. Creating high-valued data resources for research and public health requires a framework of nonconsensual access to patients' raw health data. The HIPAA Privacy Rule and the Common Rule both allow nonconsensual access to patients' data for public health and research uses.[242] If patients owned their encounter-level data, nonconsensual access for these uses still would be possible through exercise of the police and eminent domain powers.

The nub of the problem with data propertization is that it is a supply-side solution that neglects important infrastructural and demand-side issues. Access to raw patient data is necessary, but not sufficient, to ensure an adequate supply of useful data resources. Data provisioning services also are required. The prospective provision of services is inherently consensual in our system of law. The state's police and eminent domain powers only allow nonconsensual transfers of property; there is no similar mechanism to compel nonconsensual provision

241. *Id.*
242. *See* discussion *supra* Part II.B.

of services.[243] The government generally obtains services consensually, by entering into contracts,[244] requiring services in return for a grant,[245] or conditioning participation in a desirable program (such as asking hospitals to report data as a condition of Medicare eligibility).[246] The HIPAA Privacy Rule and the Common Rule have no provisions requiring nonconsensual access to data provisioning services; waivers only *permit* data holders to disclose data but do not require them to do so.[247] This is fair: data holders have only limited capacity to supply services and need discretion to refuse. Nonconsensual access to data is possible whether under a property regime or under the regulatory regime provided by the Common Rule and HIPAA Privacy Rule. Nonconsensual access to services is not possible under either regime. Access to infrastructure services, rather than the unresolved status of data ownership, is thus the key impediment to data availability.

---

243. *See* Susan W. Brenner with Leo L. Clarke, *Civilians in Cyberwarfare: Conscripts*, 43 VAND. J. TRANSNAT'L L. 1011, 1056–57 (2010) (noting, in a discussion of whether the government can require civilian information technology professionals to perform services aimed at protecting against cyberterrorism, that there has been only one instance — during the Revolutionary War — when Congress compelled civilians to provide services other than in the context of conscription for military service).

244. *See* Steven J. Kelman, *Contracting*, *in* THE TOOLS OF GOVERNMENT 282, 283–85 (Lester M. Salamon ed., 2002) (discussing features of contracts through which the government procures products or services for its use); *see also* Ruth Hoogland DeHoog & Lester M. Salamon, *Purchase-of-Service Contracting*, *in* THE TOOLS OF GOVERNMENT *supra*, 316, 320 (describing contracts in which the government procures services for delivery to third parties such as beneficiaries of welfare programs).

245. *See* David R. Beam & Timothy J. Conlan, *Grants*, *in* THE TOOLS OF GOVERNMENT, *supra* note 244, at 340, 341 (discussing the government's use of grants to stimulate performance of services).

246. *See, e.g.*, 42 C.F.R. § 482.30(c) (2010) (requiring hospitals that participate in the Medicare program to conduct utilization reviews of care provided to Medicare patients); *id.* § 482.42 (requiring hospitals that participate in the Medicare program to implement infection control programs); *id.* § 482.13(g) (requiring hospitals that participate in the Medicare program to compile statistics on deaths that occur while patients are under physical restraints and to report these statistics to the Centers for Medicare and Medicaid Services).

247. *Id.* § 164.512(i) (providing, in the HIPAA waiver provision, that uses and disclosures pursuant to a waiver are "permitted" — i.e., disclosures are allowed but not required); *id.* § 46.116(d) (couching the Common Rule's waiver provision in similarly permissive language: "An IRB may approve . . . ."). The IRB of a research institution that wishes to receive data from a data holder can approve a waiver authorizing release of the data. *See* Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,695 (Dec. 28, 2000) (rejecting, in the preamble to the HIPAA Privacy Rule, suggestions that HHS should require IRBs that approve waivers to be independent of the entity conducting the research). Under section 164.512(i), recipient-approved waivers permit the data holder to disclose data but do not require it, so the recipient has no way to force the provision of needed data and services.

IV. THE HITECH ACT'S STRATEGY FOR PROMOTING INFRASTRUCTURE DEVELOPMENT

This Part challenges claims[248] that the HITECH Act, which Congress passed as part of the 2009 economic stimulus legislation,[249] did little to promote interconnection of health information systems. It is true that the HITECH Act does not expressly require interconnection or data sharing. However, it does something arguably more important: it clarifies the price of data provisioning services, and it authorizes data holders to conduct commercial transactions for sale of those services.[250] In doing so, it lays groundwork for a commercial market in data provisioning services and provides a mechanism to finance private-sector development of health information infrastructure. The HITECH Act accepts that access to data provisioning services is inherently consensual. It authorizes a pricing structure that, if properly implemented, will create incentives for data holders and other potential service providers to "come to the market" by supplying data provisioning services within existing capacity constraints and by investing to expand capacity.

A. The Regulated Price of Infrastructure Services

At first glance, the HITECH Act purports to restrict sales of health data.[251] It states a general rule that it is unlawful for HIPAA-covered entities and their business associates to exchange a person's protected health information for direct or indirect remuneration — in other words, to sell data — unless the person authorizes the transaction.[252] However, this restriction is tempered by a list of exceptions.[253] One exception pertains to research: data holders that supply data to researchers pursuant to a HIPAA waiver[254] can charge a price that "reflects the costs of preparation and transmittal of the data."[255] In July 2010, the Office for Civil Rights ("OCR") within the HHS pro-

248. *See, e.g.*, Hall, *supra* note 23, at 635 ("[T]he economic stimulus act contains no legal requirement that funded systems actually interconnect to form a consolidated medical record for each patient . . . ."); Rodwin, *supra* note 23, at 595 (discussing the goal of "sharing of patient data for research and public uses" and noting that the "HITECH does not appear to authorize creating regulations that can achieve that goal").

249. *See* American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115.

250. *See* discussion *infra* this Part.

251. 42 U.S.C.A. § 17935(d) (West 2010 & Supp. 2011).

252. *Id.* § 17935(d)(1).

253. *Id.* § 17935(d)(2).

254. 45 C.F.R. § 164.512(i) (2010) (allowing an IRB or privacy board to waive HIPAA's usual requirement that patients authorize the use of their data in research).

255. 42 U.S.C.A. § 17935(d)(2)(B).

posed a regulation implementing this provision.[256] The proposed regulation tracks the statute closely and would permit the sale of data for use in research under a HIPAA waiver so long as the entity supplying the data receives only "a reasonable cost-based fee to cover the cost to prepare and transmit" the data.[257] The individual's permission is required only if the data supplier wishes to charge a price higher than this cost-based fee.[258]

To be clear, these provisions do not "monetiz[e] medical information."[259] The cost-based fee for data preparation and transmittal[260] is not a price for data; it is a price for data-provisioning services. Insurers, healthcare providers, and other entities that operate health databases own infrastructure, such as computer systems and software, in which they have invested to support their regular lines of business. With the aid of this infrastructure, it is possible to sift through large volumes of data, select information that meets a researcher's specifications, and process it for transmission to the researcher. The fee described in the HITECH Act[261] is for these sorts of infrastructure services. Technically speaking, the data are supplied at no charge and the fee is for services provided in responding to the data request.

The HITECH Act has another exception for public health uses of data: when supplying data for public health activities, data holders can charge a fee for data preparation and transmittal, and this fee is not subject to the cost-based cap.[262] It may at first seem wrongheaded for data holders to charge higher fees when supplying data for public health uses, which traditionally have been viewed as having a greater social value than research.[263] Yet this policy makes sense if you assume that the data supply is infrastructure-constrained. Under such

256. Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the Health Information Technology for Economic and Clinical Health Act, 75 Fed. Reg. 40,868, 40,921 (proposed July 14, 2010) (to be codified at 45 C.F.R. pts. 160, 164) (proposing a new regulation to be codified at 45 C.F.R. § 164.508(a)(4)(ii)(B), still not finalized as of this writing).

257. *Id.*

258. *See id.* (proposing new regulations to be codified at 45 C.F.R. § 164.508(a)(4)(i), 164.508(a)(4)(ii)(B), still not finalized as of this writing, requiring patient authorization before data can be disclosed for remuneration, but allowing authorization to be waived when remuneration is limited to a reasonable, cost-based fee).

259. *See* Hall, *supra* note 23, at 651 (noting that "law either prohibits monetizing medical information, or it does not clearly permit this" and proposing to allow patients to sell rights to their data).

260. 42 U.S.C.A. § 17935(d)(2)(B) (West 2010 & Supp. 2011).

261. *Id.*

262. 42 U.S.C. § 17935(d)(2)(A) (2006); *see also* Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the HITECH Act, 75 Fed. Reg. at 40,921 (proposing a new regulation to be codified at 45 C.F.R. § 164.508(a)(4)(ii)(A), still not finalized as of this writing).

263. *See* GOSTIN, *supra* note 33, at 47 (noting the high value traditionally accorded to public health activities).

conditions, a higher fee helps support investment in needed systems[264] to resolve the constraint, thus promoting wider availability of data for public health activities. By letting public health users pay more than researchers, Congress is helping ensure adequate flows of data for public health purposes. Later, when the United States has completed installation of its basic health information infrastructure, it may make sense to cap the fees for public health as well as research uses, and the HITECH Act envisions this possibility.[265]

B. Why the HITECH Act's Approach Offers Promise

In common parlance, "cost-based" means "at cost," so this new pricing scheme may not initially sound promising as a way to spur investment in interconnected data systems. However, the OCR's "reasonable, cost-based fee"[266] for data provisioning needs to be judged in light of precedents from other infrastructure industries. Health information systems are infrastructure,[267] and the HITECH Act's cost-based fee echoes cost-of-service pricing traditionally used in many other American infrastructure industries such as electric power transmission and telecommunications.[268] Historically, many of these industries exhibited natural monopoly characteristics or other structural problems that made it unwise to let prices be set by market forces.[269]

---

264. *See* CHARLES F. PHILLIPS, JR., THE REGULATION OF PUBLIC UTILITIES 172 (1993) (noting the necessity of adequate earnings to support development and expansion of infrastructure industries).

265. *See* 42 U.S.C.A. § 17935(d)(3)(B) (West 2010 & Supp. 2011) (allowing the Secretary of HHS to apply the cost-based cap on data supplied for public health use at a later time, based on an evaluation of how it would affect the availability of data). OCR is presently evaluating now whether its cost-based fee structure should apply to public health users as well as to researchers. Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the HITECH Act, 75 Fed. Reg. at 40,891.

266. Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the HITECH Act, 75 Fed. Reg. at 40,921 (proposing a new regulation to be codified at 45 C.F.R. § 164.508(a)(4)(ii)(B), still not finalized as of this writing).

267. *See* JOSÉ A. GÓMEZ-IBÁÑEZ, REGULATING INFRASTRUCTURE 4 (2003) (defining infrastructure as "networks that distribute products or services over geographical space").

268. *See* Richard A. Posner, *Natural Monopoly and its Regulation*, 21 STAN. L. REV. 548, 592 (1969) (discussing the traditional utility regulatory process, which sets regulated prices based on an "allowed cost of service [which] includes an allowance for a 'fair return' to [investors] who have provided the capital used to render the regulated service"); *see also* PHILLIPS, *supra* note 264, at 377, 381, 385 (highlighting key cost-of-service pricing principles).

269. *See* GÓMEZ-IBÁÑEZ, *supra* note 267, at 4–6 (discussing rationales for infrastructure regulation); PHILLIPS, *supra* note 264, at 51–60 (discussing natural monopoly characteristics and structural issues that may call for price regulation); Hank Intven, Jeremy Oliver & Edgardo Sepulveda, THE WORLD BANK INFORMATION FOR DEVELOPMENT PROGRAM (*INFO*DEV), TELECOMMUNICATIONS REGULATION HANDBOOK § 1.1.1 box 1-1 (Hank Intven ed., 2000) ("Where competitive markets do not exist or fail, [a widely accepted regulatory objective is to] prevent abuses of market power such as excessive pricing and anticompeti-

These concerns supplied the rationale for imposing cost-based pricing.[270] Starting with the Interstate Commerce Act of 1887, which regulated railroads,[271] Congress imposed cost-of-service pricing on the interstate shipping,[272] stockyard,[273] telephone,[274] telegraph,[275] trucking,[276] electricity,[277] natural gas,[278] and aviation[279] industries.[280] Cost-of-service pricing remained common in U.S. infrastructure regulation until late in the twentieth century, when it was partially supplanted by reforms[281] that rely more heavily on market pricing of infrastructure services.[282]

Why, at a time when cost-of-service pricing is under critique in other industries, did Congress impose a cost-based fee structure on data provisioning services? The modern critique of cost-of-service pricing focuses on its potential to be inefficient and cumbersome to administer.[283] This critique emerged late in the twentieth century, when the major policy challenge was to optimize utilization of existing infrastructures, as opposed to financing and building new infrastructures.[284] Chen notes that traditional cost-of-service infrastructure regulation may actually be the more efficient approach under economic conditions that existed earlier in the twentieth century.[285] At that time, policymakers' central challenge was to develop new infrastructures. That is the same challenge policymakers face now with respect to America's health information infrastructure — to get it built. In the

---

tive behavior . . . ."); *id.* § 5.2.2–5.2.4 (discussing market imperfections common in infrastructure industries such as telecommunications).

270. *See* PHILLIPS, *supra* note 264, at 182–83; GÓMEZ-IBÁÑEZ, *supra* note 267, at 5–6.

271. Interstate Commerce Act of 1887, Pub. L. No. 49-104, ch. 104, 24 Stat. 379, 379 (codified as amended in scattered sections of 49 U.S.C.).

272. Shipping Act of 1916, Pub. L. No. 64-260, ch. 451, 39 Stat. 728, 733–35 (codified as amended in scattered sections of 46 U.S.C.).

273. Packers and Stockyards Act of 1921, 7 U.S.C. §§ 181–229c (2006).

274. Communications Act of 1934, Pub. L. No. 73-416, ch. 652, 48 Stat. 1064, 1070 (codified as amended in scattered sections of 47 U.S.C.).

275. *Id.*

276. Motor Carrier Act of 1935, Pub. L. No. 74-255, ch. 498, 49 Stat. 543, 543 (codified as amended in scattered sections of 49 U.S.C.).

277. Public Utility Act of 1935, Pub. L. No. 74-333, ch. 687, 49 Stat. 803, 839–40 (codified as amended at scattered sections of 16 U.S.C.).

278. Natural Gas Act of 1938, 15 U.S.C. §§ 717–717w (2006).

279. Civil Aeronautics Act of 1938, ch. 601, 52 Stat. 973, 993 (repealed 1958).

280. *See* Joseph D. Kearney & Thomas W. Merrill, *The Great Transformation of Regulated Industries Law*, 98 COLUM. L. REV. 1323, 1333–34 (1998) (citing statutes imposing cost-of-service pricing on several industries).

281. *See* Jim Chen, *The Nature of the Public Utility: Infrastructure, the Market, and the Law,* 98 NW. U. L. REV. 1617, 1618 (2004) (reviewing GÓMEZ-IBÁÑEZ, *supra* note 267).

282. See Kearney & Merrill, *supra* note 280, at 1333–40.

283. *See* Chen, *supra* note 281, at 1631 (noting that public utility regulation is criticized as creating problems of "indeterminacy and inefficiency").

284. *Cf. id.* at 1620–21 (discussing the changes in infrastructure priorities from the nineteenth to the twentieth centuries).

285. *Id.* at 1633, 1650.

HITECH Act, Congress embraced a pricing structure that, during the past 100 years, has successfully financed private-sector development of many other major infrastructures in the United States.[286]

Critical to this success was the role courts played in interpreting what a "reasonable, cost-based" price must include. More than a century of Supreme Court cases have examined cost-of-service pricing in many different infrastructure contexts.[287] Under these precedents, a reasonable, cost-based fee for infrastructure services must — in order to be constitutional — let infrastructure owners recover: (1) their variable and fixed operating costs of providing services, (2) their capital investment in the infrastructure itself, and (3) a reasonable profit margin.[288] The HITECH Act's cost-based fee structure, if implemented in accordance with these precedents, would foster creation of a commercial market in the infrastructure services that are needed to convert encounter-level patient data into valuable data resources for research and public health. In July 2010, the OCR sought public comments on how, precisely, it should define the cost-based fee[289] and has not, as of this writing, issued final regulations clarifying what the fee will cover. However, the OCR presumably must heed past Supreme Court decisions that addressed cost-based pricing in other infrastructure contexts. Should the OCR fail to do so, data holders would have grounds to challenge the constitutionality of the cost-based fee. The precedents

---

286. *See supra* notes 271–280 (listing industries that were built under regulated, cost-of-service pricing).

287. *See* Written Statement, Barbara J. Evans, Law Professor, Comments on Proposed Rule RIN 0991-AB57: Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the Health Information Technology for Economic and Clinical Health Act 4-12 (2010), *available at* http://www.regulations.gov/#!documentDetail;D=HHS-OCR-2010-0016-0086 (reviewing cases in which the U.S. Supreme Court ruled on the constitutionality of cost-of-service fee structures in other infrastructure regulatory contexts)

288. *Id.*; *see, e.g.*, Bluefield Water Works & Imp. Co. v. Pub. Serv. Comm'n, 262 U.S. 679, 690 (1923) ("Rates which are not sufficient to yield a reasonable return on the value of the property . . . deprive[] the public utility company of its property in violation of the Fourteenth Amendment."); *id.* at 692–93 (identifying factors to consider in determining whether a company's allowed rate of return is confiscatory); Chicago, Milwaukee & St. Paul Ry. Co. v. Minnesota, 134 U.S. 418, 458 (1890) ("If the company is deprived of the power of charging reasonable rates for the use of its property . . . it is deprived of the lawful use of its property, and thus, in substance and effect, of the property itself, without due process of law . . . ."); *see also* PHILLIPS, *supra* note 264, at 376–82 (providing a brief history and discussion of standards the court has enunciated with respect to a fair return on invested capital); *id.* at 257–60 (reviewing Supreme Court cases that confirmed the right of utility companies to recover operating expenses, including an allowance for depreciation of invested capital).

289. Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the Health Information Technology for Economic and Clinical Health Act, 75 Fed. Reg. 40,868, 40,891 (July 14, 2010) (to be codified at 45 C.F.R. pts. 160, 164) (seeking public comment on what should be included in the cost-based fee).

strongly favor data holders' claims to receive full recovery of their operating and capital costs, plus a reasonable profit margin.[290]

Governmental intervention in markets is justified when barriers — for example, economic or legal — are blocking private-sector development of necessary infrastructure.[291] Various forms of intervention are possible, ranging from industry-specific regulation[292] to outright public ownership and operation of infrastructure.[293] The United States has rejected the latter option consistently throughout its history[294] and instead has regulated its infrastructure industries, including regulation of their pricing.[295] The HITECH Act's data sales provisions can be seen as a traditional American approach to the problem of getting major, new infrastructure developed. Rather than have the government build big databases or otherwise own health information infrastructure, the HITECH Act presumes the infrastructure will be developed, owned, and operated by the private sector subject to cost-based pricing of infrastructure services.

V. WHAT STILL NEEDS TO BE DONE

The HITECH Act's pricing provisions may improve the situation, but all is not well. There remains a widely shared perception that the HIPAA Privacy Rule and the Common Rule are blocking socially beneficial uses of data while still under-protecting individual privacy.[296] These perceptions persist, this Part argues, not because of a mere failure to propose answers; instead, the wrong questions are being asked. This Part seeks to reframe the discussion to focus on two crucial questions that fell by the wayside during the long debate[297] — from 1974 to 2002 — that produced these regulations in their current form.

---

290. *See supra* notes 287–288 and accompanying text (discussing these precedents).

291. See generally PHILLIPS, *supra* note 264, at 172–73; GÓMEZ-IBÁÑEZ, *supra* note 267, at 20–21; Chen, *supra* note 281, at 1624–28.

292. Chen, *supra* note 281, at 1628.

293. GÓMEZ-IBÁÑEZ, *supra* note 267, at 13; Daniela Klingebiel & Jeff Ruster, *Why Infrastructure Financing Facilities Often Fall Short of Their Objectives* 7 (World Bank Policy Research, Working Paper No. 2358, 2000).

294. GÓMEZ-IBÁÑEZ, *supra* note 267, at 2; Chen, *supra* note 281, at 1633 (citing STEVEN BREYER, REGULATION AND ITS REFORM 181–83 (1982)) (pointing out that the U.S. is the only nation that maintained private ownership of its major infrastructure networks, such as pipelines and power grids, throughout the entire twentieth century).

295. *See* PHILLIPS, *supra* note 264, at 171–72 (noting that rate regulation has been an important component in the regulation of public utility infrastructures).

296. *See supra* notes 3–5.

297. *See supra* notes 11–13 and accompanying text.

A. Restoring the Proper Scope of the State's Police Power to Use Data to Promote Public Health

The first neglected question is, "What is the scope of the state's police power to use private health data to promote public health?" Regulatory practice under the Common Rule conceives the scope of the state's police power more narrowly than it is conceived in other legal contexts.[298] This anomaly can be traced to an original sin during design of the Common Rule: its framers failed to define public health actions or delineate when they should be exempt from the Common Rule's consent requirements. The National Commission formed under the National Research Act of 1974[299] was instructed to delineate the boundary between research and medical treatment.[300] There was no similar directive to clarify the boundary between research and public health actions. This left a gray area in which the state's power to use data to protect the public's health is sometimes made subject to individual consent.

The Belmont Report[301] — which set the ethical principles embodied in the Common Rule — defined research as an activity that produces generalizable knowledge.[302] Using generalizability to mark the line between research and treatment worked well; it kept common "experimental" therapeutic practices, such as the off-label use of drugs in routine clinical care, from falling under the jurisdiction of the Common Rule.[303] This definition carried through to the Common Rule's definition of "human-subject research"[304] and HIPAA's definition of "research."[305] Generalizability has jurisdictional significance under the Common Rule: it delineates whether an activity is, or is not, "human-subjects research" that is regulated by the Common Rule (and thus subject to its informed consent requirements). It does not have

298. *See supra* notes 66–68 and accompanying text (discussing the state's ability to confiscate or interfere with property when the state is acting under its police power); Merrill, *supra* note 61, at 66 (characterizing legitimate exercises of police power as circumstances in which the property owner has "no entitlement").

299. National Research Act of 1974 (National Research Service Award Act of 1974), Pub. L. No. 93-348, 88 Stat. 342, 348–51 (codified as amended in scattered sections of 42 U.S.C.); *see also* HEW, 1978 REPORT, *supra* note 8, at 56,174 (publishing recommendations as required by the National Research Act of 1974).

300. U.S. DEP'T OF HEALTH, EDUC., & WELFARE, OFFICE OF THE SEC'Y, BELMONT REPORT: ETHICAL PRINCIPLES AND GUIDELINES FOR THE PROTECTION OF HUMAN SUBJECTS OF RESEARCH, 44 Fed. Reg. 23,192, 23,192 (Apr. 18, 1979) [hereinafter BELMONT REPORT] ("[T]he [National] Commission was directed to consider . . . the boundaries between biomedical and behavioral research and the accepted and routine practice of medicine . . . .").

301. *Id.*

302. *Id.* at 23,193.

303. *Id.*

304. *See* 45 C.F.R. § 46.102(d), 46.102(f) (2010) (defining "research" and "human subject").

305. *Id.* § 164.501.

similar significance under HIPAA, which creates status-based jurisdiction that depends on attributes of the data holder.[306]

The problem under the Common Rule is that generalizability of results does not provide a good bright-line rule for determining whether public health actions should or should not require consent. For example, vaccinating people to control a smallpox epidemic is permissible even without their consent;[307] vaccinating people to see which of two vaccines works better is research that obviously should require consent. Nonconsensual vaccination is justified in the first case not because it fails to produce generalizable results, but because the unvaccinated person poses a potential threat of contagion to others in the circumstances of an epidemic. Focusing on generalizability misses the point.

Ever since the Common Rule came into effect, there have been tortured efforts to draw a sensible line between "public health practice" (which does not require consent) and "public health research" (which does). Various analytical frameworks have been proposed that consider multiple factors in addition to whether generalizable knowledge is being produced.[308] The fact remains, however, that generalizability of results gives rise to a presumption that an activity is "research" that will require informed consent, and there is no clear, reproducible standard for overcoming that presumption. Public health actions that produce generalizable knowledge, with minor exceptions,[309] require informed consent.

---

306. *See id.* §§ 160.102–160.103 (defining the "covered entities" to which the HIPAA Privacy Rule applies).

307. Jacobson v. Massachusetts, 197 U.S. 11, 38–39 (1905).

308. *See, e.g.*, JAMES G. HODGE, JR. & LAWRENCE O. GOSTIN, COUNCIL OF STATE & TERRITORIAL EPIDEMIOLOGISTS, PUBLIC HEALTH PRACTICE VS. RESEARCH 7 (2004), *available at* http://www.cste.org/pdffiles/newpdffiles/CSTEPHResRptHodgeFinal.5.24.04.pdf; Amoroso & Middaugh, *supra* note 33, at 250–53; Ctrs. for Disease Control & Prevention, U.S. Dep't of Health & Human Servs., *HIPAA Privacy Rule and Public Health*, MORBIDITY & MORTALITY WKLY. REP., Apr. 11, 2003, at 1, 10, *available at* http://www.cdc.gov/mmwr/pdf/other/m2e411.pdf; James G. Hodge, Jr., *An Enhanced Approach to Distinguishing Public Health Practice and Human Subjects Research*, 33 J.L. MED. & ETHICS 125, 127–29 (2005); Dixie E. Snider, Jr. & Donna F. Stroup, *Defining Research When It Comes to Public Health*, 112 PUB. HEALTH REPS. 29, 30 (1997); CTRS. FOR DISEASE CONTROL & PREVENTION, U.S. DEP'T OF HEALTH AND HUMAN SERVS., GUIDELINES FOR DEFINING PUBLIC HEALTH RESEARCH AND NON-RESEARCH 2 (1999), *available at* http://www.cdc.gov/od/science/integrity/docs/defining-public-health-research-non-research-1999.pdf; Office for Prot. from Research Risks, *OPRR Guidance on 45 C.F.R. § 46.101(b)(5): Exemption for Research and Demonstration Projects on Public Benefit and Service Programs*, U.S. DEP'T OF HEALTH AND HUMAN SERVS., http://www.hhs.gov/ohrp/policy/exmpt-pb.html (last visited Nov. 14, 2008).

309. *See* CTRS. FOR DISEASE CONTROL & PREVENTION, U.S. DEP'T OF HEALTH AND HUMAN SERVS., GUIDELINES FOR DEFINING PUBLIC HEALTH RESEARCH AND NON-RESEARCH, *supra* note 308, at 10 (giving the example that it would be acceptable to make nonconsensual use of the health data of virus outbreak victims on a cruise ship to try to

The distinction between public health practice and public health research is extremely problematic as applied to public health uses of people's *data*, as opposed to public health actions that affect their bodies. Exercises of the state's police power can be enjoined only when they are illegitimate, as when the government acts beyond its constitutional powers or infringes a constitutional right.[310] There are real constitutional limits on the government's power to touch people's bodies.[311] Governmental touching of a person's data raises fewer constitutional problems.[312] Unconsented public health research on people's bodies would implicate constitutional protections against bodily invasion.[313] Unconsented *informational* research by a public health agency does not trigger these same concerns.

The distinction between public health practice and research, when applied to uses of people's data, has the effect of drastically narrowing the scope of the state's police power in the area of public health. The state, when legitimately exercising its police power, can require its citizens to enter nonconsensual transactions that benefit the public.[314] Any legitimate exercise of the police power — including those that produce generalizable knowledge — can support the imposition of nonconsensual requirements on citizens. Nowhere, other than under the Common Rule, does law parse legitimate exercises of the police power into those that produce generalizable knowledge — and thus require consent — and those that do not. Indeed, actions that produce generalizable knowledge offer greater benefit to the public and, if anything, present a stronger case for nonconsensual access to data.

---

identify the cause of the outbreak, even though the knowledge gained is generalizable in that it likely will benefit future cruise passengers).

310. *See* Merrill, *supra* note 61, at 70.

311. *See* Newman v. Sathyavaglswaran, 287 F.3d 786, 789 (9th Cir. 2002) ("[T]he Supreme Court repeatedly has affirmed that 'the right of every individual to the possession and control of his own person, free from all restraint or interference of others' . . . is 'so rooted in the traditions and conscience of our people,' . . . as to be ranked as one of the fundamental liberties protected by the 'substantive' component of the Due Process Clause.") (citing Schmerber v. California, 384 U.S. 757, 772 (1966) ("The integrity of an individual's person is a cherished value of our society."), Rochin v. California, 342 U.S. 165, 174 (1952) (describing unauthorized physical invasions of the body as "offensive to human dignity"), and Union Pac. Ry. Co. v. Botsford, 141 U.S. 250, 251 (1891) (discussing "the right of every individual to the possession and control of his own person, free from all restraint or interference of others")).

312. *See* Whalen v. Roe, 429 U.S. 589, 593–94, 600–04 (1977) (upholding a New York statute that let a state public health agency collect data on patients who had been prescribed Schedule II controlled substances and finding that patients had a liberty interest in informational privacy but that the interest was not fundamental under the facts of this case); Helen L. Gilbert, *Minors' Constitutional Right to Informational Privacy*, 74 U. CHI. L. REV. 1375, 1381–83 (2007) (noting significant variations in how the various federal circuits handle information privacy claims after *Whalen*, with one circuit rejecting altogether the notion that patients have a fundamental interest in informational privacy).

313. *See supra* note 311.

314. *See* Hart, *supra* note 67, at 1107.

Treating generalizability as grounds to require consent for informational research, as the Common Rule does, yields the wrong answer: consent requirements are imposed *in inverse proportion* to the amount of public benefit the use will generate.

The recent ANPRM[315] appears poised to perpetuate this problem. In a critical passage, it questions whether the Common Rule should apply to public health and quality improvement activities, but the last sentence of this passage suggests that activities that aim to produce generalizable knowledge should be regulated.[316] Successful modernization of the Common Rule requires recognition of the following point: the state's police power to protect public health encompasses a power to use data to create generalizable knowledge. The Common Rule, as currently applied, misses this point. As a result, legitimate exercises of the state's police power are being thwarted.

An example was seen recently, when Congress authorized a large health data network[317] that will use patients' clinical and insurance claims data to conduct drug safety surveillance and other studies that have the potential to produce generalizable knowledge.[318] Congress clearly has power to legislate to protect the public health,[319] and it is almost inconceivable that modern courts would question a congressional determination that the public health benefits of these activities are sufficient to warrant access to patients' data.[320] It thus seems singularly inappropriate for private IRBs to second-guess Congress's decisions. To clarify the role of IRBs in overseeing these activities, the Director of HHS's Office for Human Research Protections ("OHRP") made a determination that the congressionally authorized data uses are public health activities that are not subject to the Common Rule.[321] Nevertheless, in one recent public health study using this network, IRBs refused access to roughly five percent of the requested

315. HHS, ANPRM, *supra* note 5.

316. *Id.* at 44,521 question 24.

317. *See* 21 U.S.C.A. § 355(k)(3)(C) (West 2006 & Supp. 2011) (describing the new "[p]ostmarket [r]isk [i]dentification and [a]nalysis [s]ystem"); *see also supra* notes 203–205 and accompanying text.

318. *See* 21 U.S.C.A. § 355(k)(3)(C)(i)(I)–(VI) (West 2006 & Supp. 2011); *see* Evans, *supra* note 25, at 601–02 (discussing the purposes for which Congress authorized development of the Sentinel network).

319. *See* Parmet, *supra* note 69, at 202–03 (discussing the scope of the police power to protect public health).

320. *See* Merrill, *supra* note 61, at 63 (discussing eminent domain cases in which courts accorded "extreme deference" to legislative findings that activity offered public benefit).

321. *See, e.g.*, Letter from Jerry Menikoff, Director, Office for Human Research Prots., to Rachel E. Behrman, Acting Assoc. Dir. of Med. Policy, Center for Drug Evaluation and Research, U.S. Food & Drug Admin. (Jan 19, 2010), *in* HIPAA AND COMMON RULE COMPLIANCE IN THE MINI-SENTINEL PILOT 10, 10 (2010), *available at* http://mini-sentinel.org/work_products/About_Us/HIPAA_and_CommonRuleCompliance_in_the_Mini-SentinelPilot.pdf (deeming Sentinel activities not to be regulated by the Common Rule).

data.[322] To date, there has been a surprising lack of debate about whether it is appropriate for private IRBs to nullify congressional determinations of what is in the American public's interest.

These problems with the Common Rule could be fixed by conforming it to the HIPAA Privacy Rule's treatment of public health activities. The Privacy Rule was specifically designed to regulate disclosures and uses of data, as opposed to interventional activities, and it directly addresses public health uses of data.[323] It expressly allows data holders to make nonconsensual disclosures of data to a "public health authority that is authorized by law to collect or receive such information"[324] for various purposes, including public health "investigations"[325] — a broad term that could encompass "systematic investigation[s] . . . [that] contribute to generalizable knowledge," which is how HIPAA defines research.[326] Even if the breadth of that term is debatable, the Privacy Rule makes several things perfectly clear: the data holder does not need to conduct an IRB review[327] or make any inquiry into the nature of the intended data use when disclosing data to public health authorities.[328] It merely needs to verify that the person requesting the data is a public health official[329] with legal authority to request the data[330] and that the requested data are the minimum neces-

---

322. SARAH L. CUTRONA, ET AL., MINI-SENTINEL SYSTEMATIC VALIDATION OF HEALTH OUTCOME OF INTEREST: ACUTE MYOCARDIAL INFARCTION CASE REPORT, 10, 12 (2010), *available at* http://www.mini-sentinel.org/work_products/Validation_HealthOutcomes/Mini-Sentinel-Validation-of-AMI-Cases.pdf (noting that even when researchers requested data for a well-documented public health purpose, IRBs refused to provide 7 of 153 — or 4.6% — of the requested medical records and insisted that patient consent was required).

323. *See* 45 C.F.R. § 164.512(b) (2010) (outlining standards for disclosure and use of data for public health activities).

324. *Id.* § 164.512(b)(1)(i); *see also id.* § 164.501 (defining public health authorities to include public agencies as well as entities acting under a contract with an agency).

325. *Id.* § 164.512(b)(1)(i).

326. *Id.* § 164.501.

327. *Id.* § 164.512(b)(1); *see also* KRISTEN ROSATI ET AL., MINI SENTINEL PRIVACY PANEL, HIPAA AND COMMON RULE COMPLIANCE IN THE MINI-SENTINEL PILOT 7 (2010), *available at* http://mini-sentinel.org/work_products/About_Us/HIPAA_and_CommonRuleCompliance_in_the_Mini-SentinelPilot.pdf (stating in a White Paper published by F.D.A.'s Mini-Sentinel pilot project that "the HIPAA Privacy Rule *does not* require the covered entity [data holder] to have an IRB or Privacy Board determine whether the covered entity may make the disclosure" when disclosing protected health information to a public health authority).

328. *Id.* at 7.

329. 45 C.F.R. § 164.514(h)(2)(ii) (2010) (allowing disclosure to officials including agency employees and persons who can prove they have a contract or other authorization to act on the government's behalf).

330. *Id.* § 164.514(h)(2)(iii) (allowing the covered entity to rely on the written statement of a public agency regarding the legal authority under which it is requesting protected health information, or an oral statement if a written statement is impracticable); *see also* Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,462, 82,547 (Dec. 28, 2000) (explaining in the Preamble to the Privacy Rule that the verification process can rely on "reasonable" documentation).

sary to fulfill the public health purpose.[331] The data holder is entitled to rely on the public health authority's representations that the data request is legally authorized and meets the minimum necessary condition.[332]

These provisions differ starkly from the Common Rule's treatment of public health uses of data, but this fact is poorly understood. The IOM's 2009 study of the HIPAA Privacy Rule discussed the distinction between public health practice and public health research at some length[333] and seemed to suggest that IRBs need to parse this distinction when administering the HIPAA Privacy Rule.[334] The regulation does not require this; the HIPAA Privacy Rule envisions no IRB involvement in disclosures of data to public health authorities.[335] Misunderstandings about this fact continue to hinder public health access to data. The HIPAA Privacy Rule appropriately defers to legislatures to decide the appropriate scope of the police power to use data to protect the public's health. It is not for executive agencies, or for the private institutional review bodies they empower, to second-guess these determinations. The Common Rule makes the state's police power subject to nullification by private and potentially conflicted IRBs[336] — a matter that is all the more troubling in light of the lax procedural norms under which these bodies operate.[337]

B. Developing a Workable Doctrine of Public Use of Private Data

The second neglected question is how to determine *which* uses of data warrant nonconsensual access to private data. Existing pathways

---

331. 45 C.F.R. § 164.514(d)(3)(iii) (2010). While § 13405(b) of the HITECH Act, codified at 42 U.S.C.A. § 17935 (West 2010 & Supp. 2011), contains a provision that requires covered entities to determine what is the minimum amount of protected health information for a disclosure, recently proposed amendments to the HIPAA Privacy Rule to implement the HITECH Act did not modify a covered entity's ability to rely on minimum necessary representations by public officials. *See* Modifications to the HIPAA Privacy, Security, and Enforcement Rules Under the Health Information Technology for Economic and Clinical Health Act, 75 Fed. Reg. 40,868, 40,922–23 (July 14, 2010) (to be codified at 45 C.F.R. pts. 160, 164) (proposing a revised regulation to be codified at 45 C.F.R. § 164.514, which has not been finalized as of this writing).

332. *See supra* notes 330–31 and accompanying text.

333. IOM, PRIVACY REPORT, *supra* note 3, at 133–36.

334. *See id.* at 131, 133 (asserting that the Privacy Rule makes the same distinction between public health practice and research that the Common Rule makes and suggesting that it is important for IRBs to be able to distinguish public health practice and public health research when implementing the Privacy Rule).

335. *See supra* note 332 and accompanying text.

336. *See* Evans, *supra* note 96, at 332 (discussing potential IRB conflicts of interest); Evans, *supra* note 89, at 5 (same).

337. *See* Evans, *supra* note 96, at 332 (discussing the lax procedural requirements HIPAA imposes); Coleman, *supra* note 102, at 13–17 (discussing procedural informality of the Common Rule); Evans, *supra* note 25, at 622–25 (discussing procedural problems with the Common Rule's waiver provisions).

for nonconsensual use of data under the Common Rule and HIPAA Privacy Rule[338] were developed in an *ad hoc* manner to preserve specific uses of data that already had well-established histories before these regulations came into force. For example, health data had been widely used in research without consent in the decades before the Common Rule came into existence.[339] The regulations preserved preexisting uses without enunciating a coherent theory explaining why — and which — data uses justify nonconsensual access. The waiver provisions of the HIPAA Privacy Rule and the Common Rule lack a "public use" requirement — a criterion, similar to the one used in eminent domain jurisprudence[340] — that requires nonconsensual research uses to serve a publicly beneficial purpose.[341]

1. How the Public Use Requirement Got Lost

There is wide agreement among bioethicists that the "central ethical issue" in health informational research is to ensure that the potential public benefits are sufficient to warrant the burden on the individual.[342] At every stage of the process that led to development of the Common Rule and the HIPAA Privacy Rule, advisory bodies that pondered nonconsensual research use of data called for a utilitarian balancing of public and private interests. It is worth tracing this history because it came to an anomalous result: the waiver provisions of both regulations, as finally promulgated, lack criteria that require such a balancing.

---

338. *See* Evans, *supra* note 89, at 4 (summarizing pathways for nonconsensual use of data under the Common Rule and HIPAA Privacy Rule).

339. *See* HEW, 1978 REPORT, *supra* note 8, at 56,188 (noting that a survey of investigators, conducted as part of efforts to develop the Common Rule, found that the fact that a "study was based exclusively upon existing records" was commonly cited as a reason why consent was unnecessary).

340. *See supra* notes 70–77 and accompanying text.

341. Merrill, *supra* note 61, at 61.

342. Casarett et al., *supra* note 9, at 597 ("The central ethical issue in pharmacoepidemiologic research is deciding what kinds of projects will generate generalizable knowledge that is widely available and highly valued, and do this in a manner that protects individuals' right to privacy and confidentiality."); *see also* NATIONAL BIOETHICS ADVISORY COMMISSION, 1 ETHICAL AND POLICY ISSUES IN RESEARCH INVOLVING HUMAN PARTICIPANTS xviii, 103–04 (Aug. 2001), *available at* http://bioethics.georgetown.edu/nbac/human/overvol1.pdf (recognizing the need for nonconsensual data use in some circumstances and including, as a necessary criterion, that an IRB determine that "the benefits from the knowledge to be gained from the research study outweigh any dignitary harm associated with not seeking informed consent"); Peter D. Jacobson, *Medical Records and HIPAA: Is It Too Late to Protect Privacy?*, 86 MINN. L. REV. 1497, 1497–99 (2002) (arguing that the most important issue to resolve is which public health objectives are sufficiently important to override the individual's interest in nondisclosure).

The earliest precursor of the Common Rule was a set of 1974 regulations that required informed consent and IRB review of research, with no provision for consent to be waived.[343] The National Commission's recommendations about human-subject protections, published in 1978, focused primarily on interventional and behavioral research.[344] The Commission discussed waiving or altering consent, but not with respect to the use of data.[345] The report separately addressed research that relies on existing documents, records, or tissue specimens and stated several principles: "If the subjects are not identified or identifiable, the research need not be considered to involve human subjects," and consent requirements should not apply.[346] Even "where the subjects are identified, informed consent may be deemed unnecessary" provided certain conditions are met.[347] These conditions included a public use requirement: an IRB must determine that "the importance of the research justifies such invasion of the subjects' privacy."[348]

The Department of Health, Education, and Welfare ("HEW"), precursor of today's HHS, commenced proceedings in 1979 to incorporate the National Commission's recommendations into its existing regulations.[349] The proposed regulation did not include a waiver provision. HEW explained that it was instead considering whether certain types of behavioral research and research with data should be exempt from the regulations altogether and thus not subject to a consent requirement at all.[350] HEW sought comments on how to handle research with data. What is striking is that the unconsented use of data was, at that time, a matter of considerable indifference. Fewer than twenty commenters discussed the proposed exemption for studies with exist-

343. Protection of Human Subjects, 39 Fed. Reg. 18,914 (May 30, 1974) (codified at 45 C.F.R. pt. 46).

344. *See* HEW, 1978 REPORT, *supra* note 8.

345. *Id.* at 56,180–81 (discussing consent waivers for certain types of behavioral research).

346. *Id.* at 56,181.

347. *Id.*

348. *Id.* at 56,179; *see also id.* at 56,181 (reporting findings of a Privacy Protection Study Commission, under the auspices of the National Commission, which elaborated this balancing requirement more specifically: "[M]edical records can legitimately be used for biomedical or epidemiological research, without the individual's explicit authorization," provided that the medical care provider (who in all likelihood would have been the data holder in that era of paper records) determines "that the importance of the research or statistical purpose for which any use of disclosure is to be made is such as to warrant the risk to the individual from additional exposure of the record or information contained therein," and provided that an IRB ensures this condition has been met).

349. Proposed Regulations Amending Basic HEW Policy for Protection of Human Research Subjects, 44 Fed. Reg. 47,688 (Aug. 14, 1979) (to be codified at 45 C.F.R. pt. 46).

350. *Id.* at 47,692.

ing data,[351] whereas other issues in the proceeding drew over 500 comments.[352] Most of those who commented favored exempting research uses of data from consent requirements altogether.[353] The final rule exempted research in which the investigator records data in a de-identified manner.[354] This exemption still exists in the modern Common Rule.[355]

What HEW did not address was whether consent could be waived for research that requires access to identified or identifiable data. This type of research later gained importance as post-1980 advances in information technology made it possible to link patients' records from multiple sources to form LHRs[356] — a process that requires at least some access to identifying information.[357] The National Commission, in its 1978 report, had called for a mechanism to allow unconsented research access to identified data and records.[358] HEW and its successor, HHS, did not address this recommendation in their 1979 to 1981 rulemaking process.

The final regulation promulgated in 1981 did, however, insert a waiver provision[359] identical to the one that still exists in the Common Rule.[360] In explaining why it had inserted this provision so late in the regulatory proceedings, HHS made no reference to nonconsensual data use. Rather, the waiver provision was a response to an altogether different problem: research into the optimal design of federal benefit programs.[361] This explains why the Common Rule's waiver provision contains no public use requirement for nonconsensual data access.

---

351. Final Regulations Amending Basic HHS Policy for the Protection of Human Research Subjects, 46 Fed. Reg. 8366, 8372 (Jan. 26, 1981).

352. *Id.* at 8368.

353. *Id.* at 8372.

354. *Id.* at 8387.

355. 45 C.F.R. § 46.101(b)(4) (2010). *But see* HHS, ANPRM, *supra* note 5, at 44,518–44, 44,521, 44,527 tbl.1 (proposing to require consent for certain exempt data uses that do not presently require consent).

356. *See* discussion *supra* Part III.B.

357. *Id.*

358. HEW, 1978 REPORT, *supra* note 8, at 56,179–80.

359. Final Regulations Amending Basic HHS Policy for the Protection of Human Research Subjects, 46 Fed. Reg. 8366, 8390 (Jan. 26, 1981).

360. 45 C.F.R. § 46.116(d) (2010).

361. Final Regulations Amending Basic HHS Policy for the Protection of Human Research Subjects, 46 Fed. Reg. at 8383. HHS was responding to Crane v. Mathews, 417 F. Supp. 532 (N.D. Ga. 1976), which had held that IRB review should have applied to certain randomized studies that varied Medicaid benefits to observe impacts on beneficiaries' consumption of healthcare. HEW had responded hastily with a strained interpretation of the Common Rule that attempted to place such studies outside the scope of its regulations. *See* Secretary's Interpretation of "Subject at Risk", 41 Fed. Reg. 26,572 (Jun. 28, 1976). The issue continued to simmer and, as HHS promulgated the final revised regulations in 1981, it tried a different solution: HHS admitted that such research should be subject to IRB review but added a provision to allow waiver of informed consent. Final Regulations Amending Basic HHS Policy for the Protection of Human Research Subjects, 46 Fed. Reg. at 8383.

The waiver provision, when initially implemented, was not intended for use in approving nonconsensual data uses, so it did not incorporate the balancing test the National Commission had recommended.[362] When the waiver provision was later pressed into service for approving nonconsensual data uses, the waiver criteria were not updated for this new purpose.

The more recent HIPAA waiver provision presents a different story. The HIPAA Privacy Rule, though it has been criticized, was the product of a thoughtful and well-researched rulemaking process.[363] When developing its proposed regulation, HHS understood that waiving consent for informational research raises issues that would not be adequately addressed by simply copying the waiver criteria of the Common Rule.[364] Instead, HHS started from scratch and proposed a new set of waiver criteria. These included a requirement that an IRB make a determination that "the research is of sufficient importance so as to outweigh the intrusion of the privacy of the individual whose information is subject to the disclosure."[365] Unfortunately, this criterion drew "a large number" of adverse comments.[366] Some commenters warned that the criterion was subjective and would be inconsistently applied by IRBs; others criticized its reliance on conflicting value judgments as to whether research is important.[367]

In response to these comments, the December 2000 version of the HIPAA Privacy Rule[368] revised the balancing criterion, conforming it to a familiar test that IRBs routinely perform when approving *any* research, whether consented or unconsented: the risks of research must be reasonable in relation to the anticipated benefits of the research — if any — to the individual and the importance of the knowledge that may reasonably be expected to result from the research.[369] This change was wrongheaded. This criterion, which appears at section 46.111(a)(2) of the Common Rule, is a minimum threshold for acceptability of research.[370] Research that does not meet this criterion is considered so devoid of scientific merit that an IRB

362. *See* HEW, 1978 REPORT, *supra* note 8, at 56,181.

363. *See* Standards for Privacy of Individually Identifiable Health Information, 65 Fed. Reg. 82,462, 82,463–66 (Dec. 28, 2000) (discussing the rationale for HIPAA's various provisions in the preamble to the initial HIPAA Privacy Rule).

364. *See id.* at 82,697 (noting that the Common Rule's waiver criteria were not explicitly directed at protecting the privacy interests that the HIPAA privacy rule protects).

365. *Id.* at 82,698.

366. *Id.*

367. *Id.*

368. *Id.*

369. 45 C.F.R. § 46.111(a)(2) (2010).

370. *Id.*

cannot ethically allow people to consent to it.[371] Adopting this criterion of minimal acceptability as the criterion for approving a waiver was nonsensical: in any situation where consent could be allowed, it could be waived. This was not the sort of public use requirement the National Commission had proposed.

HHS had an opportunity to correct this error two years later, when a new administration asked HHS to revisit the HIPAA Privacy Rule.[372] Unfortunately, the correction took the form of jettisoning the troublesome balancing requirement altogether.[373] The currently effective HIPAA waiver provision, like its counterpart in the Common Rule, has no requirement that the proposed research offer any public benefit.[374] These waiver provisions are functionally equivalent to a private delegation of takings power[375] without any public use requirement whatsoever.

2. Clarifying the Concept of Public Use of Data in Research

Those who expressed concern during the first HIPAA rulemaking about IRBs' ability to balance public and private interests[376] may have had a point. Utilitarian balancing is fundamentally at odds with the autonomy-based bioethical principles these regulations seek to uphold. The interests in the balance are incommensurable.[377] Miller has pointed out that even if research has high social value, if consent is logistically difficult or impossible to obtain, and if a consent requirement may undercut the scientific validity of results, these facts "do not in themselves constitute valid ethical reasons for waiving a requirement of informed consent."[378]

The field of bioethics has drawn heavily on an atomistic concept of autonomy that portrays individuals as "self-reliant, self-governing,

---

371. *See* HEW, 1978 REPORT, *supra* note 8, at 56,180 (advancing the principle that subjects should not be exposed to research that falls below a minimal threshold of scientific quality).

372. *See* Standards for Privacy of Individually Identifiable Health Information, 67 Fed. Reg. 53,182 (Aug. 14, 2002).

373. *See id.* at 53,270.

374. *See* 45 C.F.R. § 46.116(d) (2010) (Common Rule); *id.* § 164.512(i) (2010) (HIPAA Privacy Rule).

375. *See* discussion *supra* Part II.B.

376. *See* discussion *supra* Part V.B.1.

377. For examples of the analogous critique of balancing in other contexts, see T. Alexander Aleinikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L.J. 943 (1987) (critiquing the balancing of public and private interests in twentieth-century constitutional interpretation); Richard H. Pildes, *Avoiding Balancing: The Role of Exclusionary Reasons in Constitutional Law*, 45 HASTINGS L.J. 711 (1994) (same); Jed Rubenfeld, *The First Amendment's Purpose*, 53 STAN. L. REV. 767, 778–797 (2001) (critiquing balancing as a method of First Amendment interpretation).

378. Franklin G. Miller, *Research on Medical Records Without Informed Consent*, 36 J.L. MED. & ETHICS 560, 560 (2008) (discussing but not necessarily endorsing this view).

and fundamentally alone."[379] Tauber has remarked that foundational works of modern bioethics from the years 1954 to 1970 fail to delineate how the principle of autonomy competes with other moral tenets.[380] After 1980, bioethicists began to explore alternative views of autonomy as "not merely an internal, or psychological characteristic but also an external, or social characteristic,"[381] with individuals achieving autonomy in cooperation rather than in isolation.[382] Alternatives to a consent-based model have been proposed,[383] but they lack specifics about *how* to make decisions to allow nonconsensual use of data in service of broader public interests. Modern takings jurisprudence has been equally unable to resolve such trade-offs.[384]

Nonconsensual research use of data is a "muddle"[385] strikingly similar to the one that has afflicted regulatory takings jurisprudence[386] in the years after *Penn Central Transportation Co. v. New York City*.[387] In that case, the Supreme Court applied utilitarian balancing of public and private interests to deny compensation to Penn Central when the city's Landmark Commission restricted its ability to develop the airspace above Grand Central Station, even though the restriction inflicted a major financial loss on Penn Central for the public's benefit.[388] The Court applied a deferential "rational basis" review that presumed "regulation has high social value whenever it is 'reasonably related to the promotion of the general welfare.'"[389] The HIPAA and Common Rule waiver criteria abandon even the attempt to perform utilitarian balancing. This amounts to a presumption that informational research has high social value. This arguably may be the right decision: research generates positive externalities, and it is hard to assess

379. ALFRED I. TAUBER, PATIENT AUTONOMY AND THE ETHICS OF RESPONSIBILITY 13 (2005).

380. *Id.* at 16.

381. *Id.* at 120 (quoting GRACE CLEMENT, CARE, AUTONOMY, AND JUSTICE: FEMINISM AND THE ETHIC OF CARE 22 (1996)); *see also id.* at 85 ("[I]f the self is understood as a confluence of relationships and social obligations that are constitutive of identity, then autonomy may legitimately be subordinated to other moral principles that determine how the self is governed within a social context.").

382. *Id.* at 122.

383. IOM, PRIVACY REPORT, *supra* note 3, at 254–55 (reviewing and discussing several models).

384. *See* Merrill, *supra* note 61, at 63–64 (lamenting the lack of clear standards for determining public use).

385. *See* Carol M. Rose, Mahon *Reconstructed: Why the Takings Issue Is Still a Muddle*, 57 S. CAL. L. REV. 561 (1984); Louise A. Halper, *Why the Nuisance Knot Can't Undo the Takings Muddle*, 28 IND. L. REV. 329 (1995).

386. *See* Claeys, *supra* note 58, at 1555 ("[M]odern regulatory takings law is widely recognized to be a 'muddle.'").

387. 438 U.S. 104 (1978).

388. *See id.* at 124, 130–34.

389. Claeys, *supra* note 58, at 1557 (quoting *Penn Central*, 438 U.S. at 131); *see also* Merrill, *supra* note 61, at 63–65 (discussing judicial deference to legislative findings of public benefit).

a priori which lines of research will ultimately pay off. It may be that it benefits society to encourage all research; however, this is a decision that a society needs to make through deliberation.[390] After reviewing the development of today's waiver criteria, it is obvious this deliberation never took place.

3. Developing a Workable Public Use Criterion

For waivers to merit public trust, a workable public use criterion needs to be enunciated. The takings muddle suggests by analogy that there will be no easy solution. However, it also offers a number of possible approaches that may be worth exploring.

A. Focus Not on How Decisions Should Be Made, but by Whom

Deciding which lines of research offer substantial social benefits requires a global perspective that local IRBs do not possess. A centralized, national oversight body or a legislature is better positioned to assess which lines of research warrant nonconsensual data access. One possible approach would be to form a publicly accountable body to identify general categories of research that offer public benefit. Patient advocacy groups could petition it to allow data access for research into their "pet" diseases, much as they lobby Congress for research funding for specific diseases today.[391] When Congress has authorized specific lines of health informational research, as it did in FDAAA[392] and in the comparative effectiveness provisions of the Patient Protection and Affordable Care Act of 2010,[393] this should be treated as a Blackstonian "consent of the people" to the research. Private IRBs should not be tasked with second-guessing determinations of public benefit that a duly elected legislature has already made. In an exercise of their enforcement discretion, OHRP and OCR could issue guidance creating a safe harbor that deems data holders to have

390. *See* DANIEL CALLAHAN, WHAT PRICE BETTER HEALTH? (2006) (listing issues a society should resolve through deliberation and challenging the notion that medical research is inherently good and should be pursued without regard to the burdens it places on competing values).

391. *See* REBECCA DRESSER, WHEN SCIENCE OFFERS SALVATION (2001) (discussing the role of patient advocacy groups in influencing national research policy and allocation of resources to research).

392. 21 U.S.C.A. § 355(k)(3)(C)(i)(I)-(VI) (West 2006 & Supp. 2011); *see also* Evans, *supra* note 25, at 601–02 (discussing the purposes for which Congress authorized development of the Sentinel system).

393. Pub L. No. 111-148, 124 Stat. 119 (codified as amended in scattered sections of 42 U.S.C. and I.R.C.); *see, e.g.*, *id.* § 6301; 42 U.S.C. §§ 1301–1320d-8 (2006) (amending Title XI of the Social Security Act by adding Part D — Comparative Clinical Effectiveness Research).

complied with the regulations when they make data available for legislatively authorized research uses.

B. Develop Rules of Thumb for Identifying Suspect, "Non-Public" Uses of Data

Merrill has suggested an analytical approach that focuses on identifying the traits that cause some takings to lack sufficient public purpose.[394] These presumptively private uses then can be singled out for more skeptical regulatory oversight.[395] In the same way, it may be easier to state what a publicly beneficial use of data *is not* than to specify what it *is*. The idea is to develop a list of red flags that weaken the presumption that a proposed research use of data offers public benefit. To take one example, there is not presently a health informational research registry that serves the same purpose as ClinicalTrials.gov, where sponsors of clinical trials disclose information about their planned projects.[396] It sometimes is alleged that academic and commercial researchers would be reluctant to disclose their planned informational research activities, since doing so would give away their corporate strategies and research ideas. Unwillingness to disclose research plans might be viewed as a red flag signaling that a data use has primarily a private purpose. Private-purpose data uses still would be allowed, but they would require informed consent. Persons wishing to use the public's health information must be prepared to disclose what they intend to do with it.

C. Reject Utilitarian Balancing in Favor of Natural Rights Analysis

Nineteenth-century state courts analyzed takings cases under natural rights principles that grounded property rights in personhood[397] — an approach that bears considerable resemblance to modern bioethical analysis that grounds privacy rights in autonomy. Claeys has argued rather persuasively that the old natural rights analysis did a better job of drawing sensible lines than modern utilitarian balancing can do.[398] Of particular interest are cases where state ac-

---

394. *See* Merrill, *supra* note 61, at 90–92 (identifying a need for heightened scrutiny of takings in which there is high subjective valuation of the taken property, there is a potential for secondary rent seeking, and where there has been an intentional or negligent bypass of a thick market).

395. *See id.*

396. *See* CLINICALTRIALS.GOV, http://www.clinicaltrials.gov (last visited Dec. 21, 2011).

397. Claeys, *supra* note 58, at 1577–86 (discussing nineteenth-century state courts' natural rights analysis of eminent domain cases involving state actions to abate private nuisances or to protect public health, safety, morals, and order).

398. *See id.* (examining nineteenth-century takings cases that bear similarity to regulatory takings cases and comparing them to twentieth-century regulatory takings cases); *id.* at 1556

tions force individuals to contribute positive externalities to the community: for example, by requiring homeowners to install curbs at their own expense.[399] Such cases require courts to decide whether the action is a noncompensable exercise of police power or a compensated taking.[400] This line-drawing bears conceptual similarities to the problem of distinguishing public health uses from research uses of data. In the latter problem, monetary compensation is not at stake;[401] what is at stake is whether the activity will be subject to the Common Rule's oversight requirements.

Natural rights analysis held that owners are not entitled to takings compensation when they receive "implicit in-kind" tion[402] — for example, when each homeowner who is forced to make improvements enjoys "reciprocity of advantage"[403] and will benefit from the improvements others are forced to install.[404] This was cast as the state using its police powers to force a mutually advantageous exchange that would be hard for individuals to organize by themselves; each affected person gives something to, and gets something from, the others. When there was no reciprocity of advantage — that is, when the burdens of a measure to benefit the public were disproportionately visited on some community members — the action was a taking, and compensation was owed.[405]

The notion of reciprocity of advantage survives in modern bioethical criteria for assessing whether a particular data use is "public health practice" that can be conducted without informed consent.[406] If

("Takings law gets muddled only when it applies a certain kind of utilitarian property theory to regulatory takings.").

399. *See* Palmyra v. Morton, 25 Mo. 593, 593 (1857) (upholding a town ordinance requiring homeowners to curb and pave footpaths in front of their homes at their own expense).

400. *See* Merrill, *supra* note 61, at 65 (describing a continuum of consensual transactions, compensated takings, and uncompensated confiscation or interference with property rights under the police power).

401. *See supra* notes 79–84 and accompanying text (explaining why nonconsensual research uses of data would not be compensable even if data were patient-owned).

402. Claeys, *supra* note 58, at 1589 (citing RICHARD A. EPSTEIN, TAKINGS: PRIVATE PROPERTY AND THE POWER OF EMINENT DOMAIN 195–215 (1985) and Frank I. Michelman, *supra* note 82, at 1225–26).

403. *Id.* at 1587–89, 1619–21 (tracing the "reciprocity of advantage" or "common benefit of all" concepts in nineteenth and early twentieth-century state and federal cases and noting the twentieth-century trend to supplant natural-law analysis with utilitarian principles); *id.* at 1633 (noting occasional references to reciprocity of advantage in modern Supreme Court cases but observing that modern applications have "diluted the principle so much that it is now meaningless").

404. *Id.* at 1557, 1589 (citing Paxson v. Sweet, 13 N.J.L. 196, 199 (1832)).

405. *Id.* at 1570 (discussing the early case, Vanhorne's Lessee v. Dorrance, 2 U.S. (2 Dall.) 304 (C.C.D. Pa. 1795), which enunciated the notion that there is a taking when governmental action lays "a burden upon an individual, which ought to be sustained by society at large" (*Id.* at 310)).

406. *See* HODGE & GOSTIN, *supra* note 308, at 7, 9, 52.

benefits of a study will flow primarily to the people whose data are used, as opposed to being generalizable to other populations, this weighs in favor of a finding that the study is public health practice.[407] The criterion of "benefits internal to the community" is simply reciprocity of advantage under a different name. Unfortunately, modern IRBs use this criterion in combination with other criteria — such as generalizability of results — that often muddy the waters.[408] The nineteenth-century cases treated reciprocity as central to the analysis.

A similar focus could help identify which nonconsensual uses of data are acceptable and warrant public trust even in an environment of strong respect for individual autonomy. Nonconsensual research uses of data held in large regionally or nationally scaled data networks can be conceptualized as mutually advantageous exchanges. In this light, research in very large data networks actually has stronger ethical justification than does research with smaller datasets that force "some people alone to bear public burdens which, in all fairness and justice, should be borne by the public as a whole."[409]

These and other possible solutions need to be explored. The aim here is not to advance a particular solution but rather to focus attention on the critical and long-neglected question: what is an appropriate public use of private data, and how should that decision be made?

VI. CONCLUSION

The primary aim of this Article is to critique recent data propertization proposals. While ultimately flawed, these proposals offer a kernel of conceptual value. The property analogy supplies a helpful lens for viewing a central problem in the debate about the Common Rule and HIPAA Privacy Rule — specifically, how to address the conflict between patients' desire to control their data and the public's need to use those data for various worthy purposes. This lens brings into focus two problems that have been neglected in the debate over access and privacy: that existing regulations, as applied, drastically narrow the state's police power to harness data for public benefit, and that they fail to subject nonconsensual use of data in informational research to any public use requirement. Reforms to the HIPAA Privacy Rule and the Common Rule are unlikely to resolve the debate if these problems remain unaddressed.

It is not a foregone conclusion, however, that regulatory amendments will be required to address the problems identified in this Article. The Secretary of HHS has discretion, under the existing Common

407. *Id.* at 9, 52.
408. *Id.* at 15, 47–52.
409. Armstrong v. United States, 364 U.S. 40, 49 (1960) (Harlan, J., dissenting).

Rule,[410] to determine whether its provisions apply to particular activities. Through an exercise of this authority, the Common Rule's treatment of public health uses could be brought in line with the HIPAA Privacy Rule. Guidance from the OCR could help IRBs understand their role under the latter regulation, which treats public health disclosures appropriately but is widely misunderstood. The use of guidance — by OHRP for the Common Rule, and by the OCR for the HIPAA Privacy Rule — offers a promising way to clarify criteria for granting waivers to allow nonconsensual use of data in informational research. For example, a suitably skilled and representative advisory body could grapple with the task of characterizing classes of informational research that presumptively amount to a legitimate public use. The OCR and OHRP then could exercise their enforcement discretion to issue guidance creating a safe harbor that deems IRBs to have complied with the waiver criteria when approving such uses but exposes IRBs to further scrutiny when they do not. The point of this discussion is that the problems identified in this Article are not necessarily difficult to address; they simply have been overlooked. This neglect counts as a seminal failure in federal efforts to appropriately regulate health privacy and data access over the past forty years. The long and intractable debate surrounding privacy, public health uses of data, and informational research is the open sore that came of this neglect. The neglect needs to end now.

410. 45 C.F.R. § 46.101(c) (2010).