

YALE JOURNAL OF HEALTH POLICY, LAW, AND ETHICS

VOLUME XIV

WINTER 2014

EXECUTIVE EDITORIAL BOARD

Editors-in-Chief

Sean M. Arenson
Rose Carmen Goldberg

Executive Editors

Yunsoo Kim
Julian Polaris

Managing Editor

Gwen Cassidy
Maya Hodis

SENIOR EDITORIAL BOARD

Article Editors

Peter Bozzo
Kyra Ayo Caros (Note)
Charles Felix
Brittany Ngo

Submissions Editors

Claudia Kraft
Alex Setzepfandt
Young-Hee Kim (Notes)

STAFF EDITORS

Jeremy Ader
Jessica Agatstein
Jennifer Alvisurez
Anastasia Beletsky
Nicholas Berlin
Aurelia Chaudhury
Jeffrey Chen
David Christie
Benoit Cyrenne
Gita Deo*
John Gaudet
Shobhana Gupta
Alon Gur
Monica Jordan
Habib Khan
Yunsieg Kim
Danielle Kruglak
Shefali Kumar
Rebecca Loomis
Andrea Alvarez Marin
Karl Minges

Christine H. Monahan
Raja Narayan*
Samuel Page
Lourdes Pantin
Julia Raney
Tehreem Rehman*
Marina Romani*
Ankeeta Shukla
Hanaa Siddiqi
Jacob Siegel*
Sohini Sircar
William H. Smith
Joya Sonnenfeldt
Stephen Stich*
Annabel Xulin Tan
Kimberly Miller-Tolbert
Kimberly Wachtler
Magdalena Wilson
Celeste Wong
Megan Wright*
Aileen Xu*

*Denotes Lead Editors for this issue.

ADVISORY BOARD

- Henry J. Aaron, Ph.D., *Senior Fellow, Brookings Institution*
- Lori B. Andrews, J.D., *Professor, Chicago-Kent College of Law*
- George J. Annas, J.D., M.P.H., *Professor, Boston University Schools of Law, Medicine, and Public Health*
- John D. Arras, Ph.D., *Porterfield Professor of Biomedical Ethics, University of Virginia*
- M. Gregg Bloche, M.D., J.D., *Professor of Law, Georgetown University, and Co-Director, Georgetown-Johns Hopkins Joint Program in Law & Public Health*
- Troyen A. Brennan, M.A., M.D., J.D., M.P.H., *Executive Vice President and Chief Medical Officer, CVS Caremark*
- Scott Burris, J.D., *Professor, Temple University School of Law*
- Robert A. Burt, J.D., *Alexander M. Bickel Professor of Law, Yale Law School*
- James F. Childress, Ph.D., *Professor, University of Virginia*
- Paul D. Cleary, Ph.D., *Dean and C.E.A. Winslow Professor of Epidemiology and Public Health, Yale University School of Public Health*
- Ruth R. Faden, Ph.D., M.P.H., *Philip Franklin Wagley Professor of Biomedical Ethics and Director, The Phoebe R. Berman Bioethics Institute, Johns Hopkins University*
- Lawrence O. Gostin, J.D., L.L.D., *Professor and Director, Center for Law and the Public's Health, Georgetown and Johns Hopkins Universities*
- Christine Grady, R.N., Ph.D., *Head, Section of Human Subjects Research, Department of Clinical Bioethics, National Institutes of Health*
- Dean M. Hashimoto, M.D., J.D., M.P.H., *Associate Professor of Law, Boston College, and Chief of Occupational and Environmental Medicine, Massachusetts General Hospital and the Brigham and Women's Hospital*
- Timothy S. Jost, J.D., *Robert L. Willett Family Professor of Law, Washington and Lee University School of Law*
- Ruth Katz, J.D., M.P.H., *Dean and Walter G. Ross Professor of Health Policy, George Washington University School of Public Health and Health Services*
- David Kessler, M.D., J.D., *Dean, University of California, San Francisco School of Medicine, and Former Commissioner of the U.S. Food and Drug Administration*
- Alvin K. Klevorick, Ph.D., *John Thomas Smith Professor of Law and Professor of Economics, Yale Law School and Yale University Department of Economics*
- C. Everett Koop, M.D., Sc.D., *Senior Scholar, C. Everett Koop Institute at Dartmouth College and Former Surgeon General of the United States*
- Stephen R. Latham, J.D., Ph.D., *Deputy Director, Yale's Interdisciplinary Center on Bioethics*
- Jerold R. Mande, M.P.H., *Director of Policy Programs, Yale University School of Medicine*
- Theodore R. Marmor, Ph.D., *Professor, Yale University School of Management and Yale Law School*
- Leslie Meltzer, J.D., M.Sc., *Greenwall Fellow in Bioethics and Health Policy, Georgetown University Law Center and the Johns Hopkins Bloomberg School of Public Health*
- Theodore W. Ruger, J.D., *Professor, University of Pennsylvania Law School*
- Elyn R. Saks, M.Litt., J.D., *Orrin B. Evans Professor of Law, Psychology, and Psychiatry and the Behavioral Sciences, University of Southern California Gould School of Law*

TABLE OF CONTENTS

ARTICLES

- 1 Making the Case for a Model Mental Health Advance Directive Statute**
Judy A. Clausen
- 66 Refereeing the Public Health**
Hosea H. Harvey
- 122 Human Research Subjects As Human Research Workers**
Holly Fernandez Lynch

NOTES

- 194 Criminal Law and HIV Testing: Empirical Analysis of How At-Risk Individuals Respond to the Law**
Sun Goo Lee

Making the Case for a Model Mental Health Advance Directive Statute

Judy A. Clausen*

ABSTRACT:

Acute episodes of mental illness temporarily destroy the capacity required to give informed consent and often prevent people from realizing they are sick, causing them to refuse intervention. Once a person refuses treatment, the only way to obtain care is as an involuntary patient. Even in the midst of acute episodes, many people do not meet commitment criteria because they are not likely to injure themselves or others and are still able to care for their basic needs. Left untreated, the episode will likely spiral out of control. By the time the person finally meets strict commitment criteria, devastation has already occurred. This Article argues that an individual should have the right to enter a Ulysses arrangement, a special type of mental health advance directive that authorizes a doctor to administer treatment during a future episode even if the episode causes the individual to refuse care. The Uniform Law Commissioners enacted the Uniform Health-Care Decisions Act as a model statute to address all types of advance health care planning, including planning for mental illness. However, the Act focuses on end-of-life care and fails to address many issues faced by people with mental illness. For example, the Act does not empower people to enter Ulysses arrangements and eliminates writing and witnessing requirements that protect against fraud and coercion. This Article recommends that the Uniform Law Commissioners adopt a model mental health advance directive statute that empowers people to enter Ulysses arrangements and provides safeguards against abuse. Appendix A sets forth model provisions.

* Assistant Professor, Florida Coastal School of Law. B.A., Georgetown University, 1992; J.D., University of Florida, 1995. The author would like to thank her mother, a psychiatric nurse, Judy Lacy, her brother, a psychiatrist, Dr. Benjamin Lacy, professors Heidi Anderson and Joanmarie Davoli, research assistants Dennis Uhlmann, Ashlea Edwards, Nadia Hoosien, and Timothy Wombles, and the participants of the Sixth Annual Florida Legal Scholarship Forum at Stetson University College of Law.

TABLE OF CONTENTS

INTRODUCTION 3

I. MENTAL HEALTH DIRECTIVES IN CONTEXT..... 6

 A. TYPES OF DIRECTIVES..... 6

 B. CIVIL COMMITMENT 9

 C. BENEFITS OF MENTAL HEALTH DIRECTIVES 13

 D. ADDRESSING CONCERNS ABOUT ULYSSES ARRANGEMENTS 21

II. ANALYSIS OF THE UNIFORM ACT AND STATE STATUTES 23

 A. SPECIALIZED STATUTE OR ONE SIZE FITS ALL APPROACH?..... 24

 B. REVOCATION AND ULYSSES ARRANGEMENTS 27

 C. EXECUTION 34

 D. ACTIVATION..... 38

 E. EXPIRATION..... 41

 F. ADVANCE CONSENT TO INTRUSIVE TREATMENTS, THE ROLE OF PROXIES,
 AND PATIENTS WHO FAIL TO PLAN..... 43

 G. DIRECTIVE TEMPLATES 54

III. SOLUTION: ISSUE A MODEL STATUTE 55

 A. PROVISIONS MAKING DIRECTIVES EFFECTIVE INTERVENTION TOOLS 55

 B. PROVISIONS ENSURING DIRECTIVES ARE CREATED FREE FROM UNDUE
 INFLUENCE, COERCION, OR FRAUD 56

 C. SAFEGUARDS ENSURING DIRECTIVES ARE PROPERLY IMPLEMENTED AND
 NOT ABUSED 57

 D. PROVISIONS REMOVING OBSTACLES TO ADVANCE PLANNING 58

CONCLUSION 59

**APPENDIX A: MODEL STATUTORY PROVISIONS GOVERNING MENTAL
HEALTH DIRECTIVES..... 61**

It must be remembered that for the person with severe mental illness who has no treatment the most dreaded of confinements can be the imprisonment inflicted by his own mind, which shuts reality out and subjects him to the torment of voices and images beyond our own powers to describe.¹

INTRODUCTION

The Uniform Law Commissioners (“the Commissioners”) created the Uniform Health-Care Decisions Act (“the Uniform Act”) as a comprehensive model advance directive statute for states to adopt.² The Uniform Act purports to address all types of advance health care planning, including planning for mental illness.³ However, the Commissioners focused on end-of-life decision-making, not mental illness. Therefore, the Uniform Act fails people with mental illness in several ways. Half of the states, recognizing that planning for end-of-life care implicates different issues than planning for mental health treatment, enacted separate mental health directive statutes. However, these statutes also fail to empower patients.

A key failure of the Uniform Act, which will be further discussed in Part II, is that it does not empower patients to form self-binding arrangements for care.⁴ Instead, the Uniform Act states that an individual may revoke her directive at any time.⁵ It does not expressly require capacity for revocation or allow patients to designate whether they may revoke their directives when they lack capacity. Arguably, it prevents patients from forming irrevocable directives and therefore provides no guidance on administering treatment pursuant to an irrevocable directive. For these reasons, the Uniform Act deprives patients of a valuable tool, the Ulysses arrangement.

A Ulysses arrangement is a type of mental health advance directive (mental health directive) that serves as a preventative measure for a patient to obtain treatment during an episode because the patient has learned that episodes cause

1 *Olmsted v. Zimring*, 527 U.S. 581, 609-10 (1999) (Kennedy, J., concurring in judgment).

2 Nat’l Conference of Comm’rs on Unif. State Laws, *Uniform Health-Care Decisions Act*, UNIFORM L. COMMISSION (Jan. 12, 1994), http://www.uniformlaws.org/shared/docs/health%20care%20decisions/uhcda_final_93.pdf [hereinafter UHCDA] (The Act superseded the Commissioners’ Model Health-Care Consent Act (1982), the Uniform Rights of the Terminally Ill Act (1985), and the Uniform Rights of the Terminally Ill Act (1989)). See generally Charles P. Sabatino, *The New Uniform Health Care Decisions Act: Paving a Health Care Decisions Superhighway?*, 53 MD. L. REV. 1238 (1994); Marah Stith, *The Semblance of Autonomy: Treatment of Persons with Disabilities Under the Uniform Health-Care Decisions Act*, 22 ISSUES L. & MED. 39 (2006).

3 UHCDA § 1 cmt. (stating the health care definition is to be given the broadest construction); see also Maurice S. Fisher, *Psychiatric Advance Directives and the Right to Be Presumed Competent*, 25 J. CONTEMP. HEALTH L. & POL’Y 386, 397 (2009) (asserting that UHCDA affords patients the ability to make decisions concerning future mental health issues); Sabatino, *supra* note 2, at 1240.

4 UHCDA § 3.

5 UHCDA § 3(a)–(b).

her to refuse needed intervention.⁶ The patient enters the arrangement when she has capacity.⁷ A Ulysses arrangement authorizes doctors to treat the patient during a future episode when the patient lacks capacity even if the episode causes the patient to refuse treatment at that time.⁸ A patient who enters a Ulysses arrangement essentially requests doctors to ignore the patient's illness-induced refusals.⁹

The following story illustrates why Ulysses arrangements are needed to facilitate intervention. Mr. Smith's daughter begged police to drive to her father's house and transport him to the hospital. Diagnosed with bipolar disorder, he was in the midst of an acute manic episode. His daughter had received letters from him bragging about his upcoming role in a blockbuster film. When Mr. Smith became manic, psychosis led him to a fantasy world in which he starred in a movie filmed by hidden cameras. Police agreed to check on him to determine whether he met the criteria for involuntary emergency detention and hospital admission. Hours later, police informed her that although Mr. Smith acted bizarrely, they could not transport him to the hospital against his will because his behavior did not indicate that he was a danger to himself or others.

A week later, police found Mr. Smith in front of his apartment wearing only underwear and darting into the street. Concerned that a car might hit him, the police decided that he met the criteria and transported him to the hospital against his will. Two weeks of inpatient treatment would bring Mr. Smith back to his gentle self. But 72 hours later, as required by law, doctors discharged Mr. Smith against medical advice even though he was still manic. They explained that he demanded discharge and did not meet involuntary placement criteria. Days later, police arrested Mr. Smith who was driving one hundred and twenty miles an hour on a freeway. Psychosis made him believe that he was in a televised drag race. The jail health clinic gave him lithium but failed to monitor his fluid intake. As a

6 See Elizabeth M. Gallagher, *Advance Directives for Psychiatric Care: A Theoretical and Practical Overview for Legal Professionals*, 4 PSYCHOL. PUB. POL'Y, & L. 746, 780 (1998) (providing a sample Ulysses arrangement); I. Gremmen et al., *Ulysses Arrangements in Psychiatry: A Matter of Good Care?*, 34 J. MED. ETHICS 77, 80 (2008) (preferring the term Ulysses "arrangement" over Ulysses "contract" or "statement" because "contract" overemphasizes the judicial aspects and "statement" has the connotation of a one-sided declaration); Chrisoula Andreou, *Making a Clean Break: Addiction and Ulysses Contracts*, 22 BIOETHICS 1 (2008) (arguing that there is a place for Ulysses contracts in managing addictive behavior).

7 Gremmen, *supra* note 6, at 77.

8 See Andreou, *supra* note 6, at 1. The arrangement derives its name from the main character in Homer's epic poem *Odyssey*. Ulysses was afraid the Sirens' song would lead him into danger. He directed his shipmates to tie him to the mast of his ship and not to release him, even if the Sirens' song manipulated him to demand to be set free. In the mental health context, Ulysses contracts "uphold the guidance provided by one's deepest identity conferring concerns" and potentially prevent episodes from threatening the "self". Theo Van Willigenburg & Patrick J.J. Belaere, *Protecting Autonomy as Authenticity Using Ulysses Contracts*, J. MED. & PHIL. 395, 397 (2005).

9 See Breanna M. Sheetz, Comment, *The Choice to Limit Choice: Using Psychiatric Advance Directives to Manage the Effects of Mental Illness and Support Responsibility*, 40 U. MICH. J.L. REFORM 401, 403 (2007).

result he suffered lithium toxicity, making it medically unsafe for him to take lithium. He must now rely on other treatments.

Before the onset of his illness, Mr. Smith was a mild-mannered accountant. When he takes his medication, he is still that person. Manic episodes have given him a criminal record and cost him his marriage, his career, two years of commitment in a state psychiatric hospital, and his savings. Mr. Smith wants to prevent further damage to his life by forming a Ulysses arrangement.

Acute episodes of mental illness often prevent people like Mr. Smith from realizing they are sick and cause them to refuse treatment.¹⁰ Once an episode causes a person to refuse care, the primary means of obtaining treatment is through involuntary commitment.¹¹ Even in the midst of acute episodes that have temporarily destroyed capacity, many people do not meet commitment criteria because they are not likely to injure themselves or others and are still able to care for their basic needs.¹² Requiring a person to reach a state that meets involuntary commitment criteria can postpone intervention until it is too late.¹³

This Article proposes a solution that empowers people to control their mental illnesses. Part I places mental health directives in context and begins with a description of the types of advance directives. Section I.B explains civil commitment law because a basic understanding of commitment law is necessary to appreciate the need for Ulysses arrangements. Section I.C explores the key benefit of Ulysses arrangements: to intervene early and avoid involuntary commitment. Then, it explores the benefits of mental health directives generally. Section I.D identifies and addresses concerns that Ulysses arrangements are paternalistic, create opportunities for abuse, violate due process, fail to provide contemporaneous informed consent, and destroy privacy.

Part II explores key provisions of the Uniform Act and state statutes. Section II.B compares the Uniform Act and state mental health directive statutes,

10 See KAY REDFIELD JAMISON, *AN UNQUIET MIND: A MEMOIR OF MOODS AND MADNESS* 36 (1995); Joanmarie I. Davoli, *Still Stuck in the Cuckoo's Nest: Why Do Courts Continue to Rely on Antiquated Mental Illness Research?*, 69 TENN. L. REV. 987, 1009 (2002) (asserting that inability to accept that one is mentally ill is a symptom of the disease); NAT'L ETHICS COMM., VETERANS HEALTH ADMIN., *ADVANCE DIRECTIVES FOR MENTAL HEALTH: AN ETHICAL ANALYSIS OF STATE LAWS & IMPLICATIONS FOR VHA POLICY*, NAT'L CENTER FOR ETHICS IN HEALTH CARE, U.S. DEP'T VETERANS AFF. 8 (Feb. 2008), http://www.ethics.va.gov/docs/necrpts/NEC_Report_20080220_Adv_Directives_MH-Analysis_of_State_Laws-Implications_for_VHA_Policy.pdf [hereinafter "VHA Report"] (asserting that patients entering a mania may not recognize they are sick and may refuse treatment).

11 See Sheetz, *supra* note 9, at 415.

12 See *infra* Section I.C.

13 Nick Anderson, *Dr. Jekyll's Waiver of Mr. Hyde's Right to Refuse Medical Treatment: Washington's New Law Authorizing Mental Health Care Advance Directives Needs Additional Protections*, 78 WASH. L. REV. 795, 801 (2003) (relaying testimony summarized by a legislative assistant); Davoli, *supra* note 10, at 1045 (citing Ashok K. Malla et al., *Improving Outcomes in Schizophrenia: The Case for Early Intervention*, 160 CAN. MED. ASS'N J. 843, 844 (1999) for the proposition that early intervention at the onset of psychosis improves the odds of long-term recovery).

concluding that neither adequately empowers patients to form Ulysses arrangements. Section II.C illustrates why the Uniform Act's minimal execution requirements expose patients to risks of coercion, fraud, and undue influence. Section II.D commends the Uniform Act's patient designated activation because it facilitates early intervention. Section II.E discusses when automatic expiration of mental health directives may be appropriate. Section II.F illustrates that the Uniform Act's lack of guidance on mental health treatments combined with unchecked authority to automatically selected surrogates undermines patient autonomy and potentiates abuse.

Part III explains key provisions of a model mental health directive statute set forth in Appendix A that empower patients to enter Ulysses arrangements and safeguard against abuse.

I. MENTAL HEALTH DIRECTIVES IN CONTEXT

This Part describes the status quo, which must be understood to better appreciate the need for a model statute that empowers patients to form Ulysses arrangements. Section I.A explores types of directives and explains the context in which mental health directives are implemented. Section I.B gives an overview of civil commitment law which is often the only intervention available for patients unable to form Ulysses arrangements. Section I.C explains the key benefit of a Ulysses arrangement: to intervene early and avoid commitment. Then it explains the benefits of mental health directives generally. Section I.D identifies and addresses concerns about Ulysses arrangements.

A. Types of Directives

Advance directives come in various forms.¹⁴ Instructional directives enable a patient (also known as the principal) to instruct doctors to administer care when the patient lacks capacity to provide informed consent.¹⁵ Capacity, a key concept in this Article, refers to the capacity to make and communicate health care decisions and to "understand the significant benefits, risks, and alternatives" to proposed treatment.¹⁶ Proxy directives allow a patient to appoint an agent to make health care decisions for the patient when the patient is incapacitated.¹⁷ Hybrid directives contain instructions and designate agents.¹⁸ Patients use these forms of directives for physical as well as mental illness.¹⁹ Directives intended to

14 Justine A. Dunlap, *Mental Health Advance Directives: Having One's Say*, 89 KY. L.J. 328, 347-54 (2001).

15 See John Q. La Fond & Deborah Srebnik, *The Impact of Mental Health Advance Directives on Patient Perceptions of Coercion in Civil Commitment and Treatment Decisions*, 25 INT'L J.L. & PSYCHIATRY 537-40 (2002).

16 UHCDA § 1(3).

17 102 AM. JUR. 3D *Proof of Facts* § 25 (last updated Sept. 2013).

18 La Fond & Srebnik, *supra* note 15, at 541.

19 *Proof of Facts*, *supra* note 17, at §§ 7, 25.

plan for episodes of mental illness are called mental health directives.²⁰ The Ulysses arrangement is a special type of mental health directive that is irrevocable during periods of incapacity and enables the patient to consent in advance to treatment despite illness-induced refusals.²¹ Acute episodes of mental illness often act directly to deprive patients of capacity and can distort judgment more than physical illnesses of similar severity.²² Such episodes often cause patients to refuse treatment to which they would consent if they were not influenced by the episode.²³ For this reason, clinicians implement mental health directives, including Ulysses arrangements, in a different context than they implement general advance directives (generic directives).²⁴ Professor Patricia Backlar stated that a generic directive attempts to guarantee a “good death” while a mental health directive endeavors to secure a “good life.”²⁵ Generally, doctors implement instructions regarding end-of-life treatment contained in a generic directive when the patient is in a coma.²⁶ Unlike comatose patients, a patient in the midst of an acute mental illness episode is capable of taking affirmative actions, which suggests the need for a different precautionary scenario. Doctors often must implement mental health directives during episodes in which the patient is not only conscious but unruly.²⁷ In the midst of such episodes, patients may adamantly refuse treatment requested in their directives.²⁸ Moreover, people with chronic terminal illness are more likely to receive treatment from doctors with whom they have established relationships than are people with episodic mental illness. Acute episodes of mental illness can induce people to travel, and they often receive treatment for acute episodes of mental illness in emergency rooms or, after arrest, in jails.²⁹ It is essential that mental health directives are

20 See Richard A. Van Dorn et al., *Reducing Barriers to Completing Psychiatric Advance Directives*, 35 ADMIN. & POL’Y MENTAL HEALTH 440, 448 (2008).

21 See Dunlap, *supra* note 14, at 352-55.

22 *Id.*

23 See *supra* note 10 and accompanying text.

24 Patricia Backlar, *Anticipatory Planning for Psychiatric Treatment Is Not Quite the Same as Planning for End-of-life Care*, 33 COMMUNITY MENTAL HEALTH J. 262 (1997).

25 *Id.* at 261-62.

26 See David Y. Nakashima, Comment, *Your Body, Your Choice: How Mandatory Advance Health-Care Directives Are Necessary to Protect Your Fundamental Right to Accept or Refuse Medical Treatment*, 27 U. HAW. L. REV. 201, 202-03 (2004) (discussing *In re Guardianship of Schaivo*, 780 So. 2d 176, 177 (Fla. Dist. Ct. App. 2001), in which the family of a woman in a persistent vegetative state battled over whether she should be kept alive through artificial means and stating that generic directives address situations like persistent vegetative states); Dunlap, *supra* note 14, at 356-58.

27 Robert D. Miller, *Advance Directives for Psychiatric Treatment: A View from the Trenches*, 4 PSYCHOL. PUB. POL’Y & L. 728, 734 (1998).

28 See *supra* note 10 and accompanying text.

29 See PETE EARLEY, CRAZY: A FATHER’S SEARCH THROUGH AMERICA’S MENTAL HEALTH MADNESS 2-3 (2006) (The largest public mental health facility in America is the Los Angeles County jail, which on any given day houses 3000 mentally disturbed inmates); Bureau of Justice Statistics, *Special Report: Mental Health Problems of Prison and Jail Inmates*, OFF. JUST.

enforceable wherever patients receive care.

The Uniform Act focuses on the typical end-of-life situation.³⁰ The Commissioners approved the Uniform Act because state laws for advance directives were incomplete, inconsistent, and confusing.³¹ The Commissioners are all practicing attorneys, judges, legislators, legislative staff, or law professors appointed by state governments to research, craft, and promote uniform state laws in areas where uniformity is desirable and practical.³² At the time the Uniform Act was issued, every state had one or more statutes regarding health care powers of attorney, living wills, or other forms of proxy decision-making.³³ Often, these statutes were incomplete because they only addressed a narrow set of issues or were overly formalistic and difficult for patients to follow.³⁴ The primary goals of the Uniform Act were to support patient autonomy by creating a simplified uniform process to facilitate use of advance directives and provide a method for making health care decisions when patients fail to plan.³⁵ The Commissioners simplified the formation process by dispensing with obstacles to directive formation such as requirements for a signed, witnessed, notarized writing.³⁶ The Uniform Act purported to be a comprehensive statutory scheme addressing all health care planning.³⁷ With its focus on end-of-life, the Uniform Act is not suited to the mental health context for a host of reasons explored in Part II of this paper. Most importantly, the Uniform Act does not expressly authorize patients to specify whether they may revoke their directives when they lack capacity.³⁸ It also does not expressly require capacity for revocation.³⁹ This is probably why the Uniform Act provides no guidance for administration of treatment pursuant to an irrevocable directive in the face of illness-induced refusals. For these reasons, the Uniform Act does not empower patients to form Ulysses arrangements. Without a Ulysses arrangement, a patient whose illness causes him to revoke his directive and refuse treatment has no mechanism to secure intervention unless he meets involuntary commitment criteria.

Even though state mental health directive statutes typically prohibit

PROGRAMS, U.S. DEP'T JUSTICE, <http://www.bjs.gov/content/pub/pdf/mhppji.pdf> (last updated Dec. 14, 2006) (In 2005, more than half of all inmates had a mental health problem.).

30 See *infra* Part II.

31 David M. English, *The Uniform Health-Care Decisions Act and its Progress in the States*, 15 PROB. & PROP. 19 (2001).

32 See *An Introduction to the Uniform Law Commission*, UNIFORM LAW COMM'N (July 2, 2012), <http://www.uniformlaws.org/Video.aspx?title=An%20Introduction%20to%20the%20Uniform%20Law%20Commission>.

33 Sabatino, *supra* note 2, at 1238.

34 *Id.*

35 English, *supra* note 31, at 20.

36 UHCDA § 2(a) (specifying that individual instructions “may be oral or written”).

37 English, *supra* note 31, at 20.

38 See UHCDA § 3 (describing revocation procedures, none of which allow principals to designate whether directives are revocable when the principal lacks capacity).

39 *Id.*

incapacitated patients from revoking their directives, they do not empower patients to form Ulysses arrangements. These statutes do not set forth procedures for treating patients pursuant to Ulysses arrangements. The insufficient protection that mentally ill patients receive from these statutes will be discussed at greater length in Section II.B. Without a process for administering treatment pursuant to a Ulysses arrangement, providers will discharge patients who demand discharge even if their irrevocable directives consent to treatment. This is because without a statute authorizing clinicians to administer treatment pursuant to a Ulysses arrangement, clinicians will refuse to admit and treat in the face of the patient's refusals. Doing so could expose the clinician or hospital to liability for various torts, including violating informed consent, assault, battery, false imprisonment, statutory violations, and federal civil rights violations explored in Section I.C.

B. Civil Commitment

If a patient is unable to form a Ulysses arrangement, the primary means of obtaining intervention during an episode that causes him to refuse treatment is through involuntary civil commitment.⁴⁰ The state's authority to commit people with mental illness derives from two components of sovereignty.⁴¹ The first is the police power, which is the authority to maintain peace and order.⁴² The state can confine a person who is likely to be dangerous to others.⁴³ The second is the *parens patriae* power, which enables the state to protect a person whose mental illness makes her likely to harm herself or prevents her from being able to care for her basic needs.⁴⁴

Criteria

According to the U.S. Supreme Court, civil commitment imposes a "massive curtailment of liberty"⁴⁵ that warrants strict commitment criteria. The Supreme Court decided that the clear and convincing evidence standard meets due process guarantees for civil commitment proceedings, but the preponderance of the evidence standard is inadequate.⁴⁶

For police power commitment, states typically require the government to show that because of mental illness, the person is a danger to others.⁴⁷ First, the

40 Sheetz, *supra* note 9, at 415.

41 Addington v. Texas, 441 U.S. 418, 426 (1979).

42 *Id.*

43 *Id.*

44 *Id.*

45 Zinermon v. Burch, 494 U.S. 113, 131 (1990); Vitek v. Jones, 445 U.S. 480, 491-92 (1980); Humphrey v. Cady, 405 U.S. 504, 509 (1972).

46 Addington, 441 U.S. at 432-33.

47 CHRISTOPHER SLOBOGIN ET AL., LAW AND THE MENTAL HEALTH SYSTEM: CIVIL AND CRIMINAL ASPECTS 23, 705 (5th ed. 2008).

state must prove the person suffers from a mental illness or disorder,⁴⁸ often defined as a substantial disorder of the emotional processes, thought or cognition that grossly impairs judgment, behavior or capacity to recognize reality.⁴⁹ Second, most states require proof that mental illness caused the dangerousness.⁵⁰ Third, the government must prove the dangerousness itself.⁵¹ The standards for dangerousness vary by jurisdiction. For instance, in Florida the government must show a substantial likelihood that in the near future the person will inflict serious bodily harm on another person, as evidenced by recent behavior causing, attempting, or threatening such harm.⁵² Florida is one of several states that demand a finding of an overt act as a prerequisite to involuntary civil commitment.⁵³ Not all jurisdictions have an overt act requirement.⁵⁴

For *parens patriae* commitment, states generally require the government to prove that mental illness caused the person to be a danger to herself or rendered her unable to provide for her basic needs.⁵⁵ Generally, states use the same definition of mental illness or disorder as used for police power commitment.⁵⁶ Two categories of people are potentially subject to *parens patriae* commitment. The first is people at risk of suicide or self-harm caused by provocation of others.⁵⁷ The second is people whose illnesses render them unable to provide for their basic needs for food, clothing, and shelter.⁵⁸ Typically, states that have an overt act requirement for police power commitment have the requirement for *parens patriae* commitment.⁵⁹

For both types of involuntary commitment, almost all states require consideration of less restrictive alternatives to involuntary hospitalization⁶⁰ that

48 *Foucha v. Louisiana*, 504 U.S. 71, 112 (1992); SLOBOGIN ET AL., *supra* note 47, at 723; William P. Coyle, *Cause of Action Against Psychiatrist in State-Operated Psychiatric Facility for Improper Civil Commitment*, in 10 CAUSES ACTION 2d 1, §4 (last updated Oct. 2013).

49 SLOBOGIN ET AL., *supra* note 47, at 723; *see, e.g.*, N.M. STAT. ANN. § 43-1-3(O) (2013).

50 SLOBOGIN ET AL., *supra* note 47, at 726.

51 *Id.* at 726-42 (generally addressing dangerousness); Coyle, *supra* note 48, at § 4; *see, e.g.*, *In re B.T.*, 891 A.2d 1193 (N.H. 2006) (requiring evidence of dangerous conduct and stating the psychiatrist's finding of a dangerous mental condition is insufficient for involuntary commitment).

52 FLA. STAT. § 394.467(1)(b) (2013).

53 *See, e.g.*, KAN. STAT. ANN. §59-2946(f) (2013); *Lessard v. Schmidt*, 349 F. Supp. 1078 (E.D. Wis. 1972), *vacated*, 414 U.S. 473 (1974).

54 *Matthew v. Nelson*, 461 F. Supp. 707 (N.D. Ill. 1978); MICHAEL L. PERLIN, *MENTAL DISABILITY LAW* 119 (2d ed. 2005).

55 SLOBOGIN ET AL., *supra* note 47, at 705.

56 *See supra* notes 48-49 and accompanying text.

57 DONALD H.J. HERMANN, *MENTAL HEALTH AND DISABILITY LAW IN A NUTSHELL* 159 (1997).

58 PERLIN, *supra* note 54, at 125 (This form of *parens patriae* commitment is often called "gravely disabled."); *Doe v. Gallinot*, 486 F. Supp. 983 (C.D. Cal. 1979), *aff'd*, 657 F.2d 1017 (9th Cir. 1981) (stating that the gravely disabled standard meets constitutional standards but cautioning against overbroad construction).

59 HERMANN, *supra* note 57, at 161.

60 SLOBOGIN ET AL., *supra* note 47, at 782; *see, e.g.*, FLA. STAT. § 394.467(1)(b) (2013) (prohibiting involuntary commitment without a finding that all available less restrictive treatment alternatives were adjudged inappropriate); HAW. REV. STAT. § 334-60.2 (2013).

allow for care and prevent danger, such as outpatient treatment, day or night treatment in a hospital, placement in the custody of a loved one, or home health services.⁶¹

Procedures

Emergency Detention and Screening

All states authorize involuntary emergency admission and evaluation without a full adjudicatory commitment hearing.⁶² This is the most common way a person enters the civil commitment process.⁶³ Usually, either police apprehend and transport the person to a facility or family transports the person.⁶⁴ Typically, statutes authorize police to detain and transport to a hospital a person that the officer concludes meets emergency detention and screening criteria, which are essentially the same criteria for involuntary commitment.⁶⁵ Then, a doctor at the receiving facility examines the person to determine if emergency treatment is necessary to protect the safety of the person or others.⁶⁶ States vary as to who may authorize involuntary emergency admission.⁶⁷ For example, in Virginia, only a magistrate may authorize emergency admission, but in Florida, a doctor can.⁶⁸ States also impose strict time limits under which a person may be subject to involuntary admission and examination.⁶⁹ For example, in Florida, within 72 hours from the time the person arrives at the facility, a mental health professional must examine the person to determine if she meets involuntary placement criteria.⁷⁰ If the person fails to meet the criteria, the facility must release her unless the person provides informed consent to remain as a voluntary patient.⁷¹

Involuntary Admission

Every state has formal adjudicatory procedures for involuntary commitment.⁷² Each state requires a formal commitment hearing,⁷³ with notice⁷⁴

61 *Randolph v. Cervantes*, 950 F. Supp. 771, 777 (S.D. Miss. 1996).

62 HERMANN, *supra* note 57, at 165.

63 *Id.* at 146-48.

64 Coyle, *supra* note 48, at § 2.

65 *See, e.g.*, FLA. STAT. § 394.463(2) (2013).

66 *Id.* § 394.463.

67 SLOBOGIN ET AL., *supra* note 47, at 807.

68 FLA. STAT. § 394.463 (2013); VA. CODE ANN. § 37.2-809 (2013).

69 SLOBOGIN ET AL., *supra* note 47, at 811; *see, e.g.*, VA. CODE ANN. §37.2-809 (stating the duration of temporary detention shall not exceed 48 hours before there is a hearing).

70 FLA. STAT. § 394.463(2) (2013).

71 *Id.*

72 SLOBOGIN ET AL., *supra* note 47, at 705.

73 *See, e.g.*, FLA. STAT. § 394.467(6) (2013) (hearing to be held within five days).

74 *See, e.g., ID.* § 394.4599 (2013).

and counsel,⁷⁵ and mandates periodic reviews of the legal status of committed respondents to evaluate whether they continue to meet commitment criteria.⁷⁶ In most states, a judge makes the decision to commit, but many states enable the respondent to request a jury trial.⁷⁷ Commitment hearings are often dehumanizing.⁷⁸ Many states either provide for private proceedings or allow exclusion of the respondent if being present could be harmful to the respondent.⁷⁹ Generally, states require a review hearing after initial commitment, usually from between three months to a year after admission.⁸⁰ A respondent can obtain judicial review through habeas corpus.⁸¹

Voluntary Admission

States allow for voluntary admission for inpatient mental health treatment without a hearing.⁸² According to some estimates, over half of psychiatric inpatient admissions are voluntary.⁸³ Most clinicians prefer voluntary over involuntary admission because: (1) voluntary patients are more likely to cooperate in treatment; (2) voluntary admission is less stigmatizing; and (3) involuntary commitment proceedings squander medical and judicial resources.⁸⁴ Generally, courts and legislatures also prefer voluntary treatment.⁸⁵ Critics of voluntary admission argue that admission is not truly voluntary because family and doctors frequently coerce patients to admit themselves to avoid involuntary commitment,⁸⁶ and patients often lack the capacity necessary to consent to admission.⁸⁷

75 See, e.g., MASS. GEN. LAWS ch. 123, § 5 (2013).

76 SLOBOGIN ET AL., *supra* note 47, at 705.

77 *Id.* at 820. Nebraska is one of a small number of states allowing an administrative board to commit. NEB. REV. STAT. § 71-915 (2013).

78 Karna Halverson, *Voluntary Admission and Treatment of Incompetent Persons with a Mental Illness*, 32 WM. MITCHELL L. REV. 161, 162 (2005).

79 See, e.g., FLA. STAT. § 394.467(6)(a) (2013); SLOBOGIN ET AL., *supra* note 47, at 827.

80 *Fasulo v. Arafah*, 378 A.2d 553 (1977); SLOBOGIN ET AL., *supra* note 47, at 852.

81 See, e.g., FLA. STAT. § 394.459 (2013).

82 See, e.g., *id.* § 394.4625 (2013) (providing procedures for voluntary admission and requiring discharge of voluntary patients who request discharge).

83 See Halverson, *supra* note 78, at 163 n.4.

84 *Id.* at 164.

85 Michael L. Perlin, *Competency, Deinstitutionalization, and Homelessness: a Story of Marginalization*, 28 HOUS. L. REV. 63, 116 n.305 (1991).

86 See, e.g., *In re Tiffin*, 646 N.E.2d 285 (Ill. App. Ct. 1995) (discussing a statute prohibiting statements that the patient may be subject to involuntary commitment if she does not admit herself); see also Halverson, *supra* note 78, at 166-68 (exploring arguments against voluntary admission including potential for patient coercion and lack of an adversarial process and attorney representation).

87 See, e.g., *Zinerman v. Burch*, 494 U.S. 113, 113 (1990); Perlin, *supra* note 85, at 117 (asserting that because most statutes fail to define the competency required for a valid voluntary admission, many patients who consent are incompetent).

C. Benefits of Mental Health Directives

This section first illustrates how Ulysses arrangements empower patients to intervene early and avoid commitment. Next, it enumerates the benefits of mental health directives, which include documenting informed consent (which in turn protects clinicians and facilitates treatment), improving care, safeguarding rights to refuse treatment, and avoiding guardianship.

Intervene Early and Avoid Involuntary Commitment

Forming a Ulysses arrangement is critical for some patients because it is the only effective intervention mechanism for episodes that compromise their ability to recognize their need for treatment. Involuntary commitment is the most common way patients without directives obtain intervention during an episode.⁸⁸ As explored below, Ulysses arrangements are superior to involuntary commitment because involuntary commitment comes too late and is often traumatic; the proceedings can be dehumanizing; and police apprehension can be dangerous.

The first reason early intervention through a Ulysses arrangement is better than involuntary commitment is that involuntary commitment comes too late. One patient testified, “When someone is allowed to decompose so severely before they can get help under the involuntary treatment act, they never come back quite the same.”⁸⁹ Pete Earley chronicled his struggles navigating the labyrinth of the mental health system for his son whose illness prevented him from recognizing he needed treatment:

My son was so out of control the nurse called hospital security. I was glad. Maybe now they will medicate him, I thought. But before the security guard arrived, Mike dashed outside, cursing loudly. I went after him. Meanwhile, the doctor told my ex-wife it was not illegal for someone to be mentally ill in Virginia. But it was illegal for him to treat them unless they consented. There was nothing he could do. “Even if he is psychotic?” She asked. “Yes.” Mike couldn’t forcibly be treated, the doctor elaborated, until he hurt himself or someone else.⁹⁰

The second reason intervention through a Ulysses arrangement is superior to involuntary commitment is that involuntary hospitalization in a state hospital often traumatizes patients.⁹¹ Patients may suffer symptoms of posttraumatic

⁸⁸ Sheetz, *supra* note 9, at 414.

⁸⁹ Anderson, *supra* note 13, at 801.

⁹⁰ EARLEY, *supra* note 29, at 16.

⁹¹ Cf. Letter from Thomas E. Perez, Assistant Att’y Gen., Civil Rights Div., U.S. Dep’t of Justice, to the Honorable Sonny Perdue, Ga. Office of the Governor 2 (Dec. 8, 2009), *available at* http://www.justice.gov/crt/about/spl/documents/Georgia_Psychiatric_Hospitals_findlet_12-08-

stress disorder after discharge from a state psychiatric hospital.⁹² Loved ones can only visit during limited hours.⁹³ There has been evidence of staff members verbally and even physically assaulting patients.⁹⁴ For these reasons, patients may want to consent to admission, at the first signs of an acute episode in a private hospital. Generally, patients prefer treatment from their psychiatrists whom they trust and who know their history.⁹⁵

The third reason to avoid involuntary commitment through a Ulysses arrangement is that involuntary commitment proceedings can be time-consuming, highly intrusive, and demeaning.⁹⁶ In commitment hearings, patients may witness the testimony of loved ones about the patient's behaviors during acute episodes. Patients may feel like the accused in a criminal trial.

Finally, a Ulysses arrangement potentially enables people suffering from acute episodes to avoid police apprehension, which is the typical way a person enters the commitment process.⁹⁷ For the mentally ill, police encounters can be dangerous.⁹⁸ For example, in *Drummond v. City of Anaheim*, a police encounter with Drummond, a nonviolent person with bipolar disorder and schizophrenia, resulted in officers brutally knocking Drummond unconscious, ultimately leaving Drummond in a vegetative state.⁹⁹

Document Informed Consent, Protecting Clinicians and Facilitating and Improving Care

Under modern informed consent law, physicians must provide patients relevant information about the risks and benefits of any proposed treatment and obtain the patient's informed consent before administering treatment.¹⁰⁰ Applying

09.pdf (reporting on an investigation of Georgia's state-run psychiatric hospitals). The Department found that Georgia's facilities "continue to provide deficient services that subject patients both to actual harm, and to excessive risk of serious harm, including: (1) inadequate protection from harm; (2) inappropriate mental health treatment; (3) inappropriate seclusion and restraints; [and] (4) inadequate medical care."

92 Cf. *Civil Commitment of the Mentally Ill*, 87 HARV. L. REV. 1190, 1195-97 (1974) ("[H]ospitalization itself interferes with privacy, since the patient cannot shield himself from constant observation by both his fellow patients and staff . . .").

93 See *O'Connor v. Donaldson*, 422 U.S. 563, 574-75 (1975).

94 *Civil Commitment of the Mentally Ill*, *supra* note 92, at 1197 ("Furthermore, patients in [state] hospitals risk brutality at the hands of their fellow residents and even their attendants . . .").

95 Bruce J. Winick, *Advance Directive Instruments for Those with Mental Illness*, 51 U. MIAMI L. REV. 57, 68-69 (1996).

96 *Id.*

97 See HERMANN, *supra* note 57, at 165; SLOBOGIN et al., *supra* note 47, at 806.

98 See *infra* note 99 and accompanying text.

99 *Drummond v. City of Anaheim*, 343 F.3d 1052, 1055 (2003).

100 BARRY R. FURROW ET AL., *HEALTH LAW* § 6-11 (2d ed. 2000); *Cruzan v. Director, Miss. Dep't of Health*, 497 U.S. 261, 269 (1990) (stating that the informed consent doctrine is firmly entrenched in American tort law).

the informed consent doctrine in the mental health context is problematic¹⁰¹ because, during certain phases of their illnesses, psychiatric patients may lack the capacity required to provide informed consent.¹⁰² As the U.S. Supreme Court acknowledged in *Zinerman v. Burch*, mental illness creates special problems regarding informed consent.¹⁰³ The nature of mental illness makes it foreseeable that a person needing treatment will be unable to understand forms he is being asked to sign and unable to make a knowing and voluntary decision concerning admission and treatment.¹⁰⁴ For patients with mental illness, capacity is often a fluid concept.¹⁰⁵ There are no clear legal guidelines as to what constitutes capacity.¹⁰⁶ This fluidity negatively impacts their ability to obtain treatment because doctors are rightfully concerned about administering treatment without informed consent.

When a patient with no directive becomes incapacitated, the doctor may only administer treatment by following procedures for involuntary admission and treatment or obtaining consent from a court-appointed guardian if the court has found the patient legally incompetent.¹⁰⁷ Without following such procedures, physicians who admit and treat a patient without informed consent are potentially liable for various torts, including the independent cause of action of lack of informed consent,¹⁰⁸ assault,¹⁰⁹ battery,¹¹⁰ negligence, and false imprisonment.¹¹¹ Moreover, many state mental health codes allow patients to file claims against any person who violates the patient's rights by, for example, admitting an incapacitated patient under voluntary admission procedures.¹¹² Patients of state operated facilities may also have federal civil rights claims under 42 U.S.C. § 1983 (section 1983) for due process violations if the facility admitted and treated the patient without either obtaining informed consent or following procedures for involuntary admission and treatment.¹¹³

As a record of informed consent, the directive enables the physician to admit and treat the patient during episodes when capacity is in doubt. Directives therefore protect facilities and clinicians from liability for claims based on

101 SLOBOGIN ET AL., *supra* note 47, at 290.

102 *Id.*

103 *Zinerman v. Burch*, 49 US 113, 133 (1990).

104 *Id.*

105 Halverson, *supra* note 78, at 171 (“[E]ach patient has his or her own unique mental capabilities. In addition, a patient's mental status can fluctuate in any given day, week, month or year.”).

106 *Id.* (“[T]here is no universal definition or method of determining competency.”).

107 SLOBOGIN ET AL., *supra* note 47, at 290.

108 *See, e.g., McCroskey v. University of Tennessee*, 1995 WL 329133 at *4 (Tenn. Ct. App. 1995); *Leach v. Shapiro*, 469 N.E.2d 1047, 1052 (Ohio Ct. App. 1984).

109 Coyle, *supra* note 48, at § 25.

110 *See, e.g., Allore v. Flower Hosp.*, 699 N.E.2d 560, 564 (Ohio Ct. App. 1997); *Gragg v. Calandra*, 696 N.E.2d 1282, 1286 (Ill. App. Ct. 1998).

111 Coyle, *supra* note 48, at § 24.

112 *See, e.g., FLA. STAT. § 394.459(8)(b)* (2013).

113 *See, e.g., Zinerman v. Burch*, 494 U.S. 113, 113 (1990).

admitting and treating a patient without informed consent. The directive allows the doctor to treat a psychotic patient who does not meet involuntary commitment criteria.¹¹⁴ Moreover, the Uniform Act and state mental health directive statutes provide immunity from civil or criminal liability or from discipline for unprofessional conduct for clinicians and facilities that administer treatment pursuant to a directive.¹¹⁵

*Zinerman v. Burch*¹¹⁶ illustrates how directives might protect facilities and doctors from liability. Police found Burch, bruised, bloodied, and disoriented, wandering along a highway and transported him to a private facility designated by Florida to receive mentally ill patients.¹¹⁷ Staff evaluation indicated that Burch was psychotic.¹¹⁸ In this condition, he signed forms consenting to voluntary admission.¹¹⁹ He remained in the facility for three days, was diagnosed with paranoid schizophrenia, and was administered psychotropic medication.¹²⁰ Staff determined he needed longer-term stabilization and referred him to a state psychiatric hospital.¹²¹ There, he again signed forms requesting voluntary admission and treatment even though the report of the clinician at the state hospital asserted Burch remained psychotic.¹²² Clerks simply had Burch execute voluntary admission forms, and the facility considered him a voluntary patient.¹²³ He remained at the state hospital for five months.¹²⁴ During that time, he did not have benefit of counsel.¹²⁵ No hearing was held where he could challenge his admission and treatment.¹²⁶

After discharge, Burch filed a section 1983 claim against doctors, administrators, and staff at the state hospital.¹²⁷ He alleged they deprived him of liberty without due process by admitting him as a voluntary patient when they knew or should have known he lacked capacity.¹²⁸ Florida law prohibited voluntary admission of an incapacitated patient.¹²⁹ However, Florida failed to require a capacity determination in the course of voluntary admission.¹³⁰ No one

114 See Elizabeth A. Rosenfeld, *Mental Health Advance Directives: A False Sense of Autonomy for the Nation's Aging Population*, 9 ELDER L.J. 53, 59-60 (2001).

115 UHCDA § 9(a)(3); see, e.g., Wash. Rev. Code § 71.32.170 (2013) (granting providers immunity for following a directive in good faith and without negligence).

116 494 U.S. 113 (1990).

117 *Id.* at 118.

118 *Id.*

119 *Id.*

120 *Id.*

121 *Id.* at 118-19.

122 *Id.* at 119-20.

123 *Id.* at 134.

124 *Id.* at 120.

125 *Id.* at 121.

126 *Id.* at 120.

127 *Id.* at 114-15.

128 *Id.* at 121 n.3.

129 *Id.* at 135.

130 *Id.* at 150-51.

evaluated Burch to determine whether he had capacity to provide informed consent.¹³¹ Burch argued staff should have provided him procedural safeguards required by the Due Process Clause and Florida law for involuntary admission.¹³² The Supreme Court did not rule on the merits of the section 1983 claim but held it was justiciable.¹³³ Exploration of the Court's analysis of procedural due process case law is outside the scope of this Article.¹³⁴ However, the Court's discussion sheds light on the value of mental health directives. The Court stated that even if facilities are usually justified in taking at face value a person's request for admission for medical treatment, they might not be justified in doing so without further inquiry as to a mentally ill person's request for treatment in a mental hospital.¹³⁵ Many people with mental illness lack capacity to give informed consent but do not meet commitment criteria¹³⁶ because they are not likely to injure themselves or others and are still able to care for their basic needs.

The Court discussed the involuntary commitment procedures necessary to prevent the confinement of mentally ill people who are harmless and can live safely outside the facility.¹³⁷ Involuntary confinement of these harmless individuals would violate the Constitution.¹³⁸ If Burch had had an involuntary commitment hearing, he might not have met commitment criteria.¹³⁹ A patient willing to sign voluntary admission forms but lacking capacity to provide informed consent could not be relied on to protest his voluntary admission and demand adherence to involuntary placement procedures.¹⁴⁰ Staff members were the only people able to ensure procedural protections before depriving Burch of his liberty by admitting him without his informed consent.¹⁴¹ The State may delegate to facility staff the power to admit patients, but the staff must provide constitutionally required procedural safeguards and should not escape liability when they fail to do so.¹⁴²

If Burch had a directive of which the facility was aware, the parties might have avoided litigation. If he did not want to be admitted even when he was psychotic, his directive could have made his refusal clear. Staff would have realized voluntary admission was not an option. To admit and treat Burch, they had to adhere to involuntary placement procedures. On the other hand, if Burch wanted doctors to treat him when he was psychotic, his directive could have documented his informed consent, enabling the facility to "voluntarily" admit

131 *Id.* at 113.

132 *Id.*

133 *Id.* at 149–51.

134 *Id.* at 117.

135 *Id.* at 133 n.18.

136 *Id.* at 122.

137 *Id.* at 133–34.

138 *Id.* (citing *O'Connor v. Donaldson*, 422 U.S. 563, 575 (1975)).

139 *Id.* at 134.

140 *Id.* at 135.

141 *Id.*

142 *Id.*

and treat him.¹⁴³ Modern advance directive statutes would provide immunity from civil or criminal liability for administering treatment pursuant to the directive.¹⁴⁴

Zinerman illustrates that doctors can be punished if they admit and treat patients whose capacity is in question without following an “elaborate involuntary admission process.”¹⁴⁵ As the Supreme Court acknowledged, a psychotic patient may not meet involuntary commitment criteria.¹⁴⁶ A directive is the only way patients can obtain intervention during an episode that temporarily destroys capacity. If a patient in the midst of an episode has no directive, the facility cannot voluntarily admit and treat him because he lacks capacity.¹⁴⁷ If the patient does not meet commitment criteria, even if he is psychotic, the facility cannot admit and treat him.¹⁴⁸

Directives not only document informed consent, they potentially improve treatment.¹⁴⁹ Research indicates that mental health directives provide doctors clinically useful information that can expedite and improve care.¹⁵⁰ For example, a patient who has experienced lithium toxicity may use her directive to notify doctors that administration of lithium could be dangerous. Moreover, the patient may use her directive to provide instructions for personal matters such as caring for pets in the event the patient is hospitalized. The comfort of knowing her pets will be safe may encourage the patient to voluntarily remain in the hospital for treatment until she is stable.

Safeguard Rights to Refuse Treatment

Courts have based a person’s right to refuse medical treatment on various constitutional, statutory, and common law sources. For example, in *Cruzan v. Missouri Department of Health*, the U.S. Supreme Court recognized that a competent person has a liberty interest under the Due Process Clause in refusing unwanted medical treatment, including life-sustaining hydration and nutrition.¹⁵¹ Similarly, in *Washington v. Harper*, which involved involuntary medication of a mentally ill inmate, the U.S. Supreme Court recognized a significant liberty interest under the Due Process Clause in avoiding unwanted administration of antipsychotic medication.¹⁵² Other constitutional bases support the right to refuse

143 See *supra* Section I.C.

144 E.g., HAW. REV. STAT. § 327G-10 (2013); see Sheetz, *supra* note 9, at 431.

145 *Zinerman*, 494 U.S. at 140 (O’Connor, J., dissenting); see also *supra* Section I.C.

146 *Zinerman*, 494 U.S. at 133.

147 *Id.* at 117; Perlin, *supra* note 85, at 118 (predicting that *Zinerman* would reduce voluntary admissions at state hospitals).

148 *Zinerman*, 494 U.S. at 133–34 (citing *O’Connor v. Donaldson*, 422 U.S. 563, 575 (1975)).

149 Miller, *supra* note 27, at 735–37.

150 Debra S. Srebnik et al., *The Content and Clinical Utility of Psychiatric Advance Directives*, 56 PSYCHIATRIC SERVICES 592 (2005).

151 *Cruzan v. Director, Miss. Dep’t of Health*, 497 U.S. 261, 278–79 (1990).

152 *Washington v. Harper*, 494 U.S. 210, 221 (1990).

treatment, including the “penumbral” right to privacy’s protection of bodily integrity. Moreover, the common law doctrine of informed consent,¹⁵³ state statutes,¹⁵⁴ and state constitutions support a person’s right to refuse mental health treatments.¹⁵⁵

The right to form a directive is implicit in the right to refuse treatment because a directive enables a person with capacity to prevent administration of unwanted treatment when the person lacks capacity.¹⁵⁶ *In re Rosa M.* illustrates why directives help safeguard rights to refuse treatment.¹⁵⁷ In that case, the director of a psychiatric hospital applied for an order authorizing electroconvulsive therapy (ECT) on an involuntarily committed patient. Rosa M.’s psychiatrist opined that Rosa M.’s mental illness required treatment that included ECT. However, Rosa M. lacked capacity to consent to ECT. State regulations required authorization from an immediate family member or a court order to administer ECT to a patient who lacked capacity to consent. When she had capacity, Rosa M. had executed a directive refusing ECT. This directive documented and therefore protected her right to refuse.¹⁵⁸ The court held that absent an overriding state interest, the hospital was required to honor her competent rejection of ECT even after she had lost capacity.¹⁵⁹

Avoid Guardianship

Guardianship can help people with mental illness obtain treatment during episodes that destroy the capacity necessary to provide informed consent.¹⁶⁰ The guardianship process starts when the court receives a petition to determine the incompetency of the ward and appoint a guardian.¹⁶¹ Many states allow any interested person to initiate guardianship proceedings.¹⁶² If the court determines the person is incompetent, a hearing takes place to determine whether the person needs a guardian.¹⁶³ If there is clear and convincing evidence of the need for a guardian, the court appoints one either to make all legal decisions for the ward or only specific types of decisions the ward is incompetent to make.¹⁶⁴

One advantage of a mental health directive is its potential to help the

¹⁵³ *Cruzan*, 497 U.S. at 277.

¹⁵⁴ *Id.*

¹⁵⁵ *Id.*; Michael Flaherty, *Nonconsensual Treatment of Involuntarily Committed Mentally Ill Persons with Neuroleptic or Antipsychotic Drugs as Violative of State Constitutional Guarantee*, in 74 A.L.R.4th 1099 (last updated Oct. 2013).

¹⁵⁶ Sheetz, *supra* note 9, at 423.

¹⁵⁷ *In re Rosa M.*, 597 N.Y.S.2d 544 (1991).

¹⁵⁸ *Id.* at 545.

¹⁵⁹ *Id.*

¹⁶⁰ HERMANN, *supra* note 57, at 214.

¹⁶¹ *Id.*

¹⁶² *Id.*

¹⁶³ *Id.* at 216.

¹⁶⁴ *Id.* at 216–17.

mentally ill avoid guardianship.¹⁶⁵ People with mental illness often experience long periods of full capacity and are capable of governing their lives and treatment.¹⁶⁶ Many patients do not want “any interested person”¹⁶⁷ to initiate proceedings for a judge to find them incompetent and appoint a guardian to make their decisions. Incompetency adjudications, a form of deviance labeling, can have seriously detrimental societal consequences and cause significant psychological damage to the ward.¹⁶⁸ Many patients would prefer to execute a directive in which they can appoint an agent they trust to make decisions in line with their values.¹⁶⁹ While it may be difficult for a directive to address every situation that may arise, the patient can engage in ongoing dialogue with the agent to ensure the agent understands the patient’s thoughts about treatment.¹⁷⁰ If the directive fails to address an issue, the agent can make decisions in line with the patient’s values¹⁷¹ and there will be no need for the court to appoint a guardian.¹⁷²

Some psychiatrists have asserted that guardianship poses a danger of harming the patient’s civil rights, autonomy, and independence.¹⁷³ These psychiatrists advised that doctors should recommend guardianship cautiously, only as a last resort for patients who are severely incompetent.¹⁷⁴ When the situation is less extreme, physicians should recommend other alternatives, such as mental health directives.¹⁷⁵

The process of creating a directive gives the patient a sense of empowerment and encourages self-responsibility.¹⁷⁶ The planning process is therapeutic because it provides patients opportunities to analyze the patterns of their illnesses and prevent crises.¹⁷⁷ Studies indicate patients experience a high level of satisfaction with intervention administered pursuant to a mental health directive.¹⁷⁸

165 See Winick, *supra* note 95, at 84.

166 Miller, *supra* note 27, at 731; Sheetz, *supra* note 9, at 404; see, e.g., AM. PSYCHIATRIC ASS’N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 123–39 (5th ed. 2013) [hereinafter DSM-V] (explaining that bipolar disorder is episodic); TERRI CHENEY, MANIC: A MEMOIR (2008).

167 See HERMANN, *supra* note 57, at 214.

168 Miller, *supra* note 27, at 736; Winick, *supra* note 95, at 84.

169 Winick, *supra* note 95, at 85.

170 *Id.* at 82.

171 See Dunlap, *supra* note 14, at 348 (asserting that a hybrid directive may be the most effective way to effectuate the patient’s desires).

172 See Winick, *supra* note 95, at 85.

173 Yuval Melamed et al., *Guardianship for the Severely Mentally Ill*, 19 MED. & L. 321, 325 (2000).

174 *Id.* at 325.

175 *Id.*

176 Sheetz, *supra* note 9, at 406–07.

177 Winick, *supra* note 95, at 81–82.

178 Eric B. Elbogen et al., *Effectively Implementing Psychiatric Advance Directives to Promote Self-Determination of Treatment Among People with Mental Illness*, 13 PSYCHOL. PUB. POL’Y & L. 273, 275, 285 (2007) (reporting on a study revealing that subjects reported high satisfaction with facilitated, one-on-one directive intervention).

Researchers theorize that patients with directives perceive treatment to be more self-determined because directives allow patients to actively co-author individualized mental-health crisis prevention plans.¹⁷⁹

D. Addressing Concerns About Ulysses Arrangements

Concerns

Despite the benefits of Ulysses arrangements, the concept of a self-binding directive remains controversial. Detractors argue that Ulysses arrangements are paternalistic.¹⁸⁰ This criticism is similar to concerns about *parens patriae* commitment which enables the state to intervene in a person's decisions for her benefit without regard to whether she presents a risk to others.¹⁸¹

Critics also argue Ulysses arrangements create opportunities for undue influence, abuse, and coercion by doctors and family.¹⁸² Mental illness has the potential to wreak havoc not only on the patient, but also on the patient's loved ones. There is a danger that family and treatment team members, desperate to conquer the patient's illness, will coerce the patient into forming a Ulysses arrangement that they will then use as a tool to intimidate the patient into complying with a treatment regimen.¹⁸³

Opponents argue the Ulysses arrangement violates due process because it enables a doctor to forcibly hospitalize and treat a patient even when the patient does not meet commitment criteria.¹⁸⁴ Moreover, unlike in the civil commitment context, doctors implement Ulysses arrangements without procedural protections such as an adjudicatory hearing.¹⁸⁵

Scholars also express concern about the risk of unanticipated consequences due to a patient's change of heart or failure to foresee all contingencies.¹⁸⁶ Moreover, critics argue that consent provided in a Ulysses arrangement is not valid informed consent because it is not contemporaneous.¹⁸⁷ They argue that informed consent is a continuing process in which doctors must obtain consent for each step of treatment.¹⁸⁸ When doctors treat pursuant to a Ulysses arrangement, they rely on expired consent.¹⁸⁹

¹⁷⁹ *Id.* at 274–75.

¹⁸⁰ See Rebecca S. Dresser, *Ulysses and the Psychiatrists: A Legal and Policy Analysis of the Voluntary Commitment Contract*, 16 HARV. C.R.-C.L. L. REV. 777, 785–91 (1982).

¹⁸¹ *Id.* at 785.

¹⁸² See *id.* at 852; Winick, *supra* note 95, at 94.

¹⁸³ See also Winick, *supra* note 95, at 87 (stating some people may wish to have a directive refusing hospitalization made irrevocable to prevent family from pressuring them to revoke).

¹⁸⁴ Dresser, *supra* note 180, at 800.

¹⁸⁵ *Id.* at 813–14.

¹⁸⁶ See Winick, *supra* note 95, at 88.

¹⁸⁷ Dresser, *supra* note 180, at 830–32.

¹⁸⁸ *Id.*

¹⁸⁹ *Id.*

Finally, detractors contend that to make Ulysses arrangements effective, patients must waive their privacy by notifying doctors, employers, family, and friends.¹⁹⁰

Response

A distinction should be made between different types of paternalism. Ulysses arrangements are instruments of self-paternalism, not state-paternalism, because they implement a person's rational choices instead of her illness-induced choices.¹⁹¹ This self-paternalism respects patient autonomy by empowering the patient to direct her health care even during episodes that destroy the capacity necessary to provide informed consent.¹⁹² Depriving patients of the right to form Ulysses arrangements is itself a form of state-paternalism because it presumes to decide for the patient what is best for her.

It is true that doctors implement Ulysses arrangements without the procedural protections provided in civil commitment. However, the liberty deprivation involved in implementing a Ulysses arrangement is minimal compared to the deprivation of freedom involved in involuntary commitment, as discussed in Section I.C. When doctors implement a Ulysses arrangement, they follow a patient's advance written instructions.¹⁹³ Involuntarily committed patients do not provide advance consent. Because implementation of a Ulysses arrangement involves hospitalizing a patient despite her contemporaneous objections, the enabling statute should provide procedural protections.¹⁹⁴ The model provisions provide such protections because they: (1) limit self-binding hospitalization to three weeks,¹⁹⁵ (2) require doctors to heed treatment refusals from patients with capacity, (3) require express written consent before administering psychotropic medication in contravention of illness-induced objections, (4) prohibit Ulysses arrangements for ECT and psychosurgery, and (5) allow patients to seek injunctive relief.

While Ulysses arrangements may create opportunities for abuse, the potential of the significantly more coercive environments of civil commitment and incarceration outweighs concerns of coercion from family and doctors. To protect against abuse, the enabling statute should impose safeguards to ensure patients form Ulysses arrangements voluntarily and doctors implement Ulysses arrangements in strict compliance with patient instructions. A patient with capacity should always be able to revoke her Ulysses arrangement.¹⁹⁶ This

190 *Id.* at 851.

191 *Id.*

192 See Willigenburg and Belaere, *supra* note 8, at 395–96.

193 Roberto Cuca, *Ulysses in Minnesota: First Steps Toward a Self-Binding Psychiatric Advance Directive Statute*, 78 CORNELL L. REV. 1152, 1153 (1993).

194 *Id.* at 1154, 1182–85.

195 See Dresser, *supra* note 180, at 781 n.15 (stating that three weeks is the suggested length).

196 VHA Report, *supra* note 10, at 9.

Article's model provisions impose such safeguards including requirements for: (1) capacity determinations at the time of directive formation and implementation, (2) witnessing by multiple disinterested people, and (3) review and approval by two psychiatrists before administering treatment pursuant to a Ulysses arrangement.

To address concerns about unanticipated contingencies, patients and doctors should engage in ongoing dialogue and update the directive to ensure it remains current.¹⁹⁷ The directive might also use broad language to enable doctors to administer the most therapeutic treatment. For example, a directive could consent to psychiatric medication generally, rather than specifying the particular medication. Like this Article's model, the enabling statute should grant patients the right to designate an agent who can implement the patient's wishes when unforeseen contingencies occur. Moreover, the model provisions provide for automatic expiration of Ulysses arrangements every two years. This mechanism helps ensure the Ulysses arrangement continues to represent the patient's wishes.

Although informed consent provided through a Ulysses arrangement is admittedly not truly contemporaneous, if patients are unable to form Ulysses arrangements, they become victims of their illnesses. When an episode causes them to refuse treatment, they cannot obtain intervention until they are deemed dangerous as defined by the state commitment statute.¹⁹⁸ The Ulysses arrangement empowers the patient to determine her care when she lacks capacity to provide informed consent.¹⁹⁹ Informed consent in the directive is valid because Ulysses arrangements are only appropriate for patients who have already experienced previous episodes and responded to treatment.²⁰⁰ Mental illnesses often follow a pattern that enables the patient and the doctor to predict the intervention necessary to address future episodes.²⁰¹ Moreover, automatic expiration of Ulysses arrangements ensures the patient provided consent relatively recently. This will be discussed in Section II.E.

The criticism that Ulysses arrangements destroy privacy, like the others, presumes to weigh the benefits and risks of Ulysses arrangements for the patient. Acute episodes damage patients' lives and health. Patients should be able to decide that preventing future humiliating psychotic episodes that compromise their privacy justifies disclosing the Ulysses arrangement to their inner circles.

II. ANALYSIS OF THE UNIFORM ACT AND STATE STATUTES

Part II explains key provisions of the Uniform Act and state advance

197 Winick, *supra* note 95, at 81–86.

198 Cuca, *supra* note 193, at 1152.

199 *Id.* at 1153.

200 *Id.*; Dresser, *supra* note 180, at 800–01 (stating that confinement based on a Ulysses arrangement would allow for treatment that had been successful for the patient's past episodes).

201 *Id.* at 847–51 (asserting that it would not respect patient liberty to permit enforcement of Ulysses arrangements where there is a high likelihood of predictive error).

directive statutes. It illustrates that the Uniform Act, with its focus on end-of-life, ignores the needs of people with mental illness. For example, the Uniform Act: (1) does not empower patients to form Ulysses arrangements; (2) dispenses with execution requirements which protect against abuse; (3) fails to provide guidance on advance consent to common mental health treatments; (4) develops an automatic surrogate selection system which exposes mental health patients to undue risks; and (5) provides a template which is inappropriate in the mental health context. In addition, this Part explains how key provisions of state advance directive statutes fail to empower people with mental illness.

A. Specialized Statute or One Size Fits All Approach?

Before critiquing the Uniform Act, it is necessary to describe its primary purpose and a few of its key provisions. The Uniform Act, approved in 1993, strives to pave a health care decision-making superhighway.²⁰² The doctrine of informed consent gives people with capacity the right to determine their treatment.²⁰³ When patients lack capacity, they need a mechanism through which to exercise control over their care. The Uniform Act provides this mechanism. It allows an individual with capacity to give oral or written instructions to a provider.²⁰⁴ These instructions remain in force even after the person loses capacity.²⁰⁵ The patient may also execute a written power of attorney for health care that authorizes an agent to make health care decisions when the patient lacks capacity.²⁰⁶ Moreover, a patient may orally designate a surrogate decision-maker.²⁰⁷ For patients who fail to plan, the Uniform Act sets forth a system for automatically selecting a surrogate.²⁰⁸ The surrogate is bound by the patient's instructions or known wishes.²⁰⁹ When there are none, the surrogate must act in the patient's best interests.²¹⁰

Only nine states²¹¹ have adopted the Uniform Act, likely because most states do not want to revisit their existing advance health care planning legislation just to make small improvements.²¹² Probably because mental illness planning implicates different issues than end-of-life planning, half of the states enacted

202 Sabatino, *supra* note 2, at 1238.

203 See *supra* Part I.C.

204 UHCDA § 2.

205 *Id.* § 2(b), 7(d).

206 *Id.* § 2.

207 *Id.* § 5(b).

208 *Id.*

209 *Id.* §§ 2(e), 5(f).

210 *Id.* §§ 2(e), 5(f).

211 See ALA. STAT. § 22-8A-1 to 22-8A-13 (2013); ALASKA STAT. § 13.52.010–52.395 (2013); CAL. PROB. CODE §§ 4670–4743 (2013); DEL. CODE ANN. tit. 16, §§ 2501–2518 (2013); HAW. REV. STAT. § 327E-1 to -16 (2013); ME. REV. STAT. tit. 18-A, §§ 5-801 to 5-817 (2013); MISS. CODE ANN. § 41-41-201 to -229 (2013); N.M. STAT. ANN. § 24-7A-1 to -18 (2013); WYO. STAT. ANN. § 35-22-401 to 35-22-416 (2013); see also Stith, *supra* note 2, at 40 n.1.

212 English, *supra* note 31, at 19.

separate mental health directive statutes.²¹³ Generic directive statutes govern mental health directives in the remaining states.²¹⁴ The Uniform Act is a model generic directive statute because it is not a specialized mental health directive statute. Instead, it purports to govern all advance health care planning, for both physical and mental illness.²¹⁵ Generic directive statutes focus on end-of-life issues. For example, Florida has no separate mental health directive statute.²¹⁶ The legislative findings of Florida's generic directive statute address end-of-life and palliative care but fail to mention psychotropic medication or ECT.²¹⁷ Similarly, the recommended statutory forms in Illinois, Wisconsin, and Alaska for health care instructions make the instructions effective when the physician determines the patient has a terminal condition or is in a persistent vegetative state.²¹⁸ This provision is inappropriate for patients who need their mental health directives to take effect when they lose capacity due to an acute episode. Moreover, the provision is potentially confusing and upsetting for otherwise healthy patients with mental illness. Having to confront emotionally charged end-of-life issues at the same time as a mental health patient plans for the next episode might be disturbing for an already vulnerable patient.

The Uniform Act and state generic directive statutes created for end-of-life situations fail to address issues people with mental illness frequently face. This void burdens patients and hospitals with unnecessary litigation. In *Cohen v. Bolduc*,²¹⁹ litigation escalated to the level of the state Supreme Court. Because Massachusetts had no mental health directive statute, the Massachusetts Supreme Court had to address whether the state general health care proxy statute authorized an agent to commit a principal to a mental health facility.²²⁰ The patient's proxy was activated in the summer of 2000.²²¹ The Massachusetts Supreme Court did not issue its decision until January of 2002.²²² Undoubtedly, the patient, her family, and her hospital wasted time and incurred unnecessary expense and emotional strain when the parties had to bring their case all the way to the state Supreme Court to resolve an avoidable situation. Of course, enacting a mental health directive statute will not obviate litigation. However, when mental health questions arise, courts will have the benefit of legislative

213 See *Cohen v. Bolduc*, 760 N.E.2d 714, 719 (Mass. 2002); VHA Report, *supra* note 10, at 3, 14–15 (listing Arizona, Hawaii, Idaho, Illinois, Indiana, Kentucky, Louisiana, Maine, Maryland, Michigan, Minnesota, Montana, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Dakota, Tennessee, Texas, Utah, Washington, and Wyoming).

214 See VHA Report, *supra* note 10, at 3.

215 UHCDA § 1(5).

216 FLA. STAT. §§ 765.101–.205 (2013).

217 *Id.* § 765.102 (2013).

218 See VHA Report, *supra* note 10, at 3; ALASKA STAT. § 13.52.010 (2013); 755 ILL. COMP. STAT. §43/5 (2013); WIS. STAT. § 155.05 (2013).

219 760 N.E.2d 714, 719 (Mass. 2002).

220 *Id.* at 715.

221 *Id.* at 716.

222 *Id.* at 714.

guidance.²²³

When a court applies a generic directive statute in a mental health crisis, the court is in the untenable position of interpreting a law intended to address end-of-life care, not mental illness. For instance, in the *Cohen* case, the court had to decide whether commitment authority was implicit in the generic directive statutory scheme.²²⁴ The health care proxy statute defined “health care” to include any treatment, service or procedure to diagnose or treat the patient’s physical or mental condition.²²⁵ From this reference to mental conditions, the court decided that the legislature did not intend to limit the agent’s authority.²²⁶ As *Cohen* acknowledged, the legislature had never addressed the commitment issue.²²⁷ The Court was forced to survey other states’ mental health directive statutes, which were split as to whether an agent possessed commitment authority.²²⁸ *Cohen* looked to the Uniform Act, which only allows an agent or surrogate to commit if the principal expressly provided commitment authority in a written directive.²²⁹ Other state legislatures, whose statutory definitions of health care included treatment of mental conditions, decided that an agent did not have commitment authority.²³⁰

In a legislative vacuum, the *Cohen* court made a policy decision that express commitment authority is not required for an agent to commit the principal.²³¹ *Cohen* made this policy decision despite the fact that several state advance directive statutes prohibit an agent from committing a principal.²³² Reasonable lawmakers disagree on the issue.²³³ *Cohen* underscores the need for a model mental health directive statute for state elected officials to adopt, enabling elected lawmakers to give guidance on key mental health issues. The process of enacting a mental health directive statute requires the legislature to address issues implicated in advance planning for acute episodes. When the legislature considers the proposed legislation, it will review testimony from experts and stakeholders, including psychiatrists and patients. This process will enable the legislature to develop sound policy for the state.²³⁴

223 See *infra* Part III.

224 *Cohen*, 760 N.E.2d at 719–20.

225 *Id.* at 720.

226 *Id.*

227 *Id.* at 718.

228 *Id.* at 718–19.

229 *Id.* at 718 n.14.

230 *Id.* at 718 n.15.

231 *Id.* at 721.

232 *Id.* at 718.

233 *Id.*

234 See, e.g., Anderson, *supra* note 13, at 801 (exploring patient testimony before the Washington legislature when it considered the bill).

B. Revocation and Ulysses Arrangements

Requirements for a Ulysses Enabling Statute

To empower patients to form Ulysses arrangements, the enabling statute must have a few key components. First, it must enable patients to choose to form a directive that is irrevocable during periods of incapacity. A patient who cannot form an irrevocable directive cannot enter a Ulysses arrangement. Because episodes often cause patients to refuse treatment and revoke their directives, irrevocable directives are necessary to enable patients to secure treatment despite contemporaneous refusals.²³⁵

However, allowing patients to form irrevocable directives does not, in and of itself, empower patients to form Ulysses arrangements. The enabling statute must set forth procedures for administering treatment in the face of contemporaneous objections.²³⁶ Without a well-defined process and clear authority, a doctor typically will not force treatment on a refusing patient based only on the fact that her directive is irrevocable when she lacks capacity. Even with the typical statutory statement of provider immunity, doctors will be rightfully concerned about liability for unlawfully administering involuntary treatment.

Physician reluctance to treat in the face of contemporaneous objections is not the only concern. A more serious concern involves risks of coercion, undue influence, and fraud when doctors forcibly hospitalize and treat a patient even when the patient does not meet commitment criteria.²³⁷ The enabling statute must provide procedural protections to ensure that patients form Ulysses arrangements knowingly and voluntarily and that doctors implement the arrangements in strict accordance with patient instructions.²³⁸

The Uniform Act Approach

The Uniform Act does not empower patients to form Ulysses arrangements for several reasons. First, it does not allow patients to choose whether they can revoke their directives when they lack capacity.²³⁹ The revocation provision allows an individual to revoke the designation of an agent by a signed writing or by notifying her physician.²⁴⁰ Individuals may revoke at any time and in any manner that communicates intent to revoke portions of the directive that do not designate a surrogate, such as instructions.²⁴¹ The revocation provision does not

²³⁵ Cuca, *supra* note 193, at 1173.

²³⁶ *Id.* at 1181–85.

²³⁷ Dresser, *supra* note 180, at 800.

²³⁸ Cuca, *supra* note 193, at 1154.

²³⁹ UHCDA § 3, 11.

²⁴⁰ *Id.* § 3(a).

²⁴¹ *Id.* § 3(b).

expressly require capacity to revoke.²⁴² However, the Uniform Act capacity provision states “an individual is presumed to have capacity to make a health-care decision, to give *or revoke* an advance health-care directive, and to designate or disqualify a surrogate.”²⁴³ The commentary explains that this is a rebuttable presumption.²⁴⁴ These provisions are subject to two different interpretations, neither of which empowers patients to form Ulysses arrangements.

Under the first interpretation, the Uniform Act precludes patients from forming irrevocable directives. This is because it does not expressly require capacity to revoke and does not allow patients to designate whether their directives are revocable during periods of incapacity.²⁴⁵ This might be why some states that implemented the Uniform Act amended the Act’s language to expressly require capacity for revocation.²⁴⁶ For example, the New Mexico statute states that “an individual *while having capacity* may revoke.”²⁴⁷ The New Mexico legislature recognized that if it wanted to require capacity for revocation, an amendment was necessary. If the Commissioners intended to require capacity for revocation or to give patients the choice, the Uniform Act would have done so expressly.

Comparison to state statutes that allow patients to choose supports the conclusion that the Uniform Act does not give patients the choice. For example, the Arizona statute states that “*unless limited by the express authority in this document, a principal even if incapable, may revoke*” her mental health directive.²⁴⁸ The Uniform Act does not contain such language.²⁴⁹ In other instances in which the Uniform Act provides patients a choice, it does so expressly. For example, the Uniform Act states that a directive becomes active when the patient lacks capacity *unless the patient provides otherwise*.²⁵⁰ The Uniform Act revocation provision does not contain such language.²⁵¹ Therefore, pursuant to this first interpretation, the Uniform Act precludes patients from forming irrevocable directives.

However, there is another possible interpretation. Under this second interpretation, the Uniform Act requires capacity to revoke a directive. This alternative interpretation relies on the Uniform Act’s rebuttable presumption of capacity to revoke the directive.²⁵² Arguably, the rebuttable capacity presumption implies that *only* patients with capacity may revoke. If someone is able to rebut

242 See *id.* § 3.

243 *Id.* § 11(b) (emphasis added).

244 *Id.* § 11, cmt.

245 *Id.* §§ 3, 11.

246 ME. REV. STAT. tit. 18-A, §5-803 (2013); N.M. STAT. §24-7A-3 (2013); WYO. STAT. ANN. §35-22-404 (2013).

247 N.M. STAT. § 24-7A-3 (emphasis added).

248 ARIZ. REV. STAT. § 36-3285 (2013) (emphasis added).

249 Cf. UHCDA § 3.

250 *Id.* § 2(c).

251 Cf. *id.* § 3.

252 *Id.* § 11 cmt.

the capacity presumption, the incapacitated patient will be prevented from revoking her directive. Under this interpretation, at least in theory, patients may form irrevocable directives.

Even assuming the Uniform Act requires capacity for revocation, it does not empower patients to form Ulysses arrangements for two reasons. First, the Commissioners did not describe how to rebut the capacity presumption.²⁵³ This omission may be why some states that implemented the Uniform Act added instructions.²⁵⁴ For example, Maine added the following statement: “This presumption may be rebutted by a determination by the individual’s primary physician or by a court of competent jurisdiction.”²⁵⁵ Without guidance like this, it is unlikely anyone will try, much less succeed in rebutting the capacity presumption. The patient’s revocation will stand even if it was induced by an episode. This is especially true because one purpose of the Uniform Act is to place health care decisions in the hands of the patient, the family, and providers, not the courts.²⁵⁶ Rebutting the capacity presumption would likely involve court intervention which the Commissioners sought to avoid.²⁵⁷ Even assuming a Uniform Act capacity requirement for revocation, the Commissioners’ failure to provide guidance on rebutting the capacity presumption prevents patients from forming self-binding arrangements.

Even if one interprets the Uniform Act to require capacity to revoke, the Uniform Act does not empower people to form Ulysses arrangements for a second reason. It provides no process for administering treatment pursuant to an irrevocable directive in the face of patient refusals. It does not set forth clear authority and procedural protections necessary to overcome physician reluctance to treat in the face of illness-induced refusals and to protect patients from coercion and abuse.

States’ Approaches

A majority of all states allow the principal to revoke a generic directive at any time, even if she has lost capacity.²⁵⁸ However, most states with mental health directive statutes only allow a patient with capacity to revoke a mental

²⁵³ See *id.*

²⁵⁴ See, e.g., ME. REV. STAT. tit.18-A, §5-811 (2013); N.M. STAT. §24-7A-11 (2013); WYO. STAT. ANN. §35-22-412 (2013).

²⁵⁵ ME. REV. STAT. tit. 18-A, § 5-811.

²⁵⁶ See BARRY FURROW ET AL., HEALTH LAW, CASES, MATERIALS, AND PROBLEMS 814 (7th ed. abr. 2013).

²⁵⁷ See *supra* note 254 and accompanying text; UHCDA §6 cmt. (stating that courts have no particular expertise with respect to healthcare decision-making and court involvement causes delays).

²⁵⁸ See VHA Report, *supra* note 10, at 9 (noting that in 36 out of 50 states, incapacitated patients may revoke a generic directive).

health directive.²⁵⁹

Patients should not be forced to form irrevocable directives. The majority of states with separate mental health directive statutes reinforce the stigma of mental illness when they do not allow patients to choose whether to make their directives revocable during periods of incapacity.²⁶⁰ The rationale for prohibiting incapacitated patients from revoking mental health directives is that preferences articulated in a written directive more likely reflect the authentic values of the patient than choices made when the patient is incapacitated.²⁶¹ The rationale is based on the premise that restricted revocation during periods of incapacity best serves patient autonomy because it respects a patient's choices made when she was able to thoroughly consider the risks and benefits of treatment options.²⁶² This premise applies equally for generic directives as it does for mental health directives. Therefore, there is no policy reason to restrict revocation of mental health directives if the state does not restrict revocation of generic directives, and most states do not.²⁶³

Restricted revocation only for mental health directives undermines parity for mental health care. In one study, almost half of surveyed patients indicated they wanted authority to revoke their mental health directives during periods of incapacity.²⁶⁴ In states that do not allow incapacitated patients to revoke their mental health directives, patients only have the power to create self-binding directives.²⁶⁵ They cannot create individualized mental health care plans. One of the Commissioners' goals was to encourage patients to form advance directives.²⁶⁶ Patients have greater autonomy if they provide advance instructions and designate agents who understand their preferences.²⁶⁷ Patients who want the power to revoke their mental health directives when they lack capacity may refrain from advance planning if their only option is to form an irrevocable directive.

Typical mental health directive statutes do not empower patients to form Ulysses arrangements because they fail to set forth a process for administering treatment in the face of illness-induced refusals.²⁶⁸ The typical statute merely requires capacity for revocation and provides immunity to physicians who follow

259 *Id.* (noting that 18 out of 25 states with mental health directive statutes only allow revocation from patients with capacity).

260 *See id.* at 9; Cuca, *supra* note 193, at 1162–63; Gallagher, *supra* note 6, at 778; Srebnik et al., *supra* note 150, at 592.

261 VHA Report, *supra* note 10, at 9; Cuca, *supra* note 193, at 1162–63.

262 VHA Report, *supra* note 10, at 9.

263 Gallagher, *supra* note 6, at 778; VHA Report, *supra* note 10, at 9.

264 Srebnik et al., *supra* note 150, at 592.

265 *See* Dresser, *supra* note 180, at 781; Sabatino, *supra* note 2.

266 Sabatino, *supra* note 2, at 1238–39.

267 *Id.* at 1239.

268 Dresser, *supra* note 180, at 781 (Several safeguards would have to be in place for a Ulysses arrangement.).

the directive.²⁶⁹ This is insufficient to enable physicians to forcibly hospitalize and treat patients who do not meet commitment criteria. Physicians will be legitimately fearful of liability for administering involuntary treatment. A clear process with safeguards against abuse helps address concerns that family and providers will coerce patients into forming Ulysses arrangements that they will use to force treatment on the patient. Patients should only enter Ulysses arrangements voluntarily and knowingly. Therefore, such arrangements are not appropriate for patients deprived of the right to choose whether they can revoke their directives when they lack capacity.

Washington's Approach

Washington has a unique approach that provides instructions on implementation of an irrevocable directive. When a principal's mental health directive remains irrevocable during incapacity and consents to inpatient mental health treatment, but the principal refuses admission, the facility may admit the patient despite illness-induced refusals.²⁷⁰ There are strict criteria for such admission.²⁷¹ First, one doctor in conjunction with another²⁷² must determine whether the principal lacks capacity. The Washington statute does not address whether a principal's refusal of admission in contravention of express instructions in her directive supports a determination of incapacity.²⁷³ This is a failure because many mental illnesses can induce people to refuse admission and treatment.²⁷⁴ The directive provides clear evidence that the patient, when she had capacity, requested admission and treatment. Second, the doctor must obtain the informed consent of the principal's agent if one is designated.²⁷⁵ Third, after evaluation, the doctor must determine and make a written finding that the principal needs inpatient evaluation or treatment that cannot be accomplished in a less restrictive setting.²⁷⁶ Fourth, the doctor must document in the patient's medical record a summary of findings and recommendations.²⁷⁷

If the doctor determines the principal has capacity, the principal may only be

269 See, e.g., HAW. REV. STAT. § 327G-4 (2013) (merely stating that capacity is required to revoke); 755 ILL. COMP. STAT. § 43/5 (2013); Sheetz, *supra* note 9, at 431 (Many statutes grant immunity for providers who make good faith efforts to comply with directives.).

270 WASH. REV. CODE § 71.32.140 (2013).

271 *Id.* § 71.32.140(2).

272 *Id.* § 71.32.140(2)(a)–(3). The statute requires a physician or psychiatric registered nurse practitioner, in conjunction with another health care provider, to make the incapacity determination. If the admitting clinician is not a psychiatrist/psychiatric advanced registered nurse practitioner, a mental health professional shall assess the principal within 24 hours to determine continued need for inpatient evaluation or treatment.

273 *Id.* § 71.32.140.

274 See *supra* note 10 and accompanying text.

275 WASH. REV. CODE § 71.32.140(2)(b).

276 *Id.* § 71.32.140(2)(c).

277 *Id.* § 71.32.140(2)(d).

admitted or remain in inpatient treatment if the principal consents or is detained under involuntary commitment law.²⁷⁸ If two doctors determine that the principal lacks capacity and the principal continues to refuse admission, the principal may seek injunctive relief.²⁷⁹ The facility may retain the patient for up to 14 days, and only for the amount of time that she consented to inpatient treatment in her directive. At that point, the facility must discharge the patient unless she regains capacity and consents to further treatment or is detained under involuntary commitment law.²⁸⁰

The incapacitated principal's instructions in her directive control her treatment with one significant exception.²⁸¹ Even if the principal's irrevocable directive consents to inpatient treatment despite illness-induced refusals, the facility shall discharge the principal if she "takes actions demonstrating a desire to be discharged, in addition to making statements requesting to be discharged."²⁸² The facility shall not use restraint in any way to prevent discharge.²⁸³ This limitation essentially prevents patients from entering Ulysses arrangements.²⁸⁴

Even the Washington approach, noted for its progressive support of patient empowerment,²⁸⁵ falls short of authorizing Ulysses arrangements. First, it requires a facility to discharge an incapacitated patient who takes action and makes statements demonstrating the desire to be discharged, even if discharge contravenes the patient's irrevocable directive.²⁸⁶ The following illustrates why this prevents Ulysses arrangements.

A patient executes an irrevocable directive consenting to inpatient treatment that becomes active pursuant to its terms. His daughter drives him to a hospital where he refuses admission. The admitting psychiatrist follows the Washington protocol and determines that the patient lacks capacity and needs the inpatient treatment his directive describes.²⁸⁷ The mentally ill patient does not recognize that he is ill. He demands discharge through words and actions. Psychiatrists determine that although the patient lacks capacity, he fails to meet involuntary commitment criteria. Left untreated, his mental illness will likely escalate to

278 *Id.* § 71.32.140(4)(a).

279 *Id.* § 71.32.140(4)(d).

280 *Id.* § 71.32.140(5).

281 *Id.* § 71.32.140(6)(b).

282 *Id.*

283 *Id.*

284 *See id.* (providing that because this is a voluntary admission, a patient who takes action to leave and demands discharge must be discharged unless she meets involuntary commitment criteria but failing to explain why the patient's illness-induced demands override her consent in her directive).

285 *See* Sheetz, *supra* note 9, at 401 (stating that Washington authorizes Ulysses directives and advocating for other states to adopt similar provisions).

286 WASH. REV. CODE § 71.32.140(6)(b).

287 *Id.* § 71.32.140(2).

psychosis.²⁸⁸ Despite this inevitability, Washington requires discharge.²⁸⁹

Second, Washington fails to assist doctors in their assessment of patient capacity when a principal's illness-induced refusals contradict the patient's directive.²⁹⁰ When a principal arrives at a facility but refuses admission because of an episode, Washington requires a capacity assessment.²⁹¹ If doctors determine the principal lacks capacity, they may admit the principal only if they follow strict protocols.²⁹² If the principal has capacity, doctors must discharge the principal unless the principal consents to inpatient treatment.²⁹³ The Washington statute fails to recognize that a person who refuses care requested in her irrevocable directive necessarily exhibits substantial evidence of incapacity.

A person cannot have capacity if he does not understand the significant benefits of proposed treatment. Acute episodes can destroy insight and cause patients to refuse intervention.²⁹⁴ When a patient's irrevocable directive consents to treatment that the patient refuses when he arrives at the hospital, the refusal itself is evidence of incapacity. In this way, the Washington statute ignores several factors that make doctors reluctant to admit patients whose illnesses cause them to refuse treatment. First, only a small percentage of patients with mental illness execute mental health directives.²⁹⁵ Therefore, most psychiatrists have little experience implementing directives generally, much less Ulysses arrangements. Second, psychiatrists are very familiar with the strict criteria for involuntary admission and treatment. Unless Washington instructs otherwise, doctors will likely automatically apply this strict criteria. Third, because capacity is fluid, capacity determinations are not black and white decisions. When in doubt, doctors will likely err on the side of caution and discharge patients whose illnesses cause treatment refusals regardless of consent to treatment in an irrevocable directive.²⁹⁶ This caution prevents necessary intervention. Honoring a patient's consent to early intervention in an irrevocable directive not only respects patient autonomy, it could potentially save the patient's life.

288 See *supra* note 13 and accompanying text.

289 WASH. REV. CODE § 71.32.140(6)(b).

290 *Id.* § 71.32.140; see also *id.* § 71.32.110.

291 *Id.* § 71.32.140(2).

292 *Id.*

293 *Id.* § 71.32.140(4)(a).

294 See *supra* note 10 and accompanying text.

295 See Maria J. O'Connell & Catherine H. Stein, *Psychiatric Advance Directives: Perspectives of Community Stakeholders*, 32 ADMIN. & POL'Y MENTAL HEALTH 241, 244 (2005) (Only 6.8% of people with schizophrenia surveyed had a mental health directive.); Jeffrey Swanson et al., *Psychiatric Advance Directives Among Public Mental Health Consumers in Five U.S. Cities: Prevalence, Demand, and Correlates*, 34 J. AM. ACAD. PSYCHIATRY & L. 43, 54 (2006) (finding that between 4 and 13% of mental health patients surveyed had directives).

296 See *supra* notes 107–113 and accompanying text.

C. Execution

The Uniform Act's Minimal Execution Requirements

Under the Uniform Act, an "individual instruction" is the principal's directions about her health care.²⁹⁷ Oral instructions are valid.²⁹⁸ The patient's physician need only record the oral instructions in the principal's medical record.²⁹⁹ A patient may issue written instructions without any witnesses, notarization, or mandatory form or language.³⁰⁰

The Uniform Act permits three types of proxies to make health care decisions for patients who lack capacity: surrogates, guardians, and agents.³⁰¹ A principal's designation of an agent (a power of attorney for health care) must be in a signed writing.³⁰² The only people who may not be agents are owners, operators, or employees of residential long-term health care institutions where the principal receives care, unless they are related to the principal.³⁰³ Designation of an agent need not be witnessed or notarized.³⁰⁴ A surrogate is an individual authorized to make the principal's health care decisions when the principal lacks capacity, and no agent or guardian has been designated or is available.³⁰⁵ A patient may select a surrogate orally by personally informing her doctor.³⁰⁶

A comparison with state advance directive statutes is useful to illustrate that many legislatures consider execution requirements to be useful protections against abuse. Every state that implemented the Uniform Act imposed witnessing requirements for all directives, presumably to protect against fraud and coercion.³⁰⁷ The need to protect against coercion in the context of mental health directives is arguably greater than the need for generic directives. Scholars and legislatures have recognized the potential risk of family and doctors using mental health directives as instruments to coerce patients to accept certain treatments.³⁰⁸ Patients with mental illness are especially vulnerable to coercion because they may perceive the threat of involuntary commitment or forced administration of

297 UHCDA § 1(9).

298 *Id.* § 2(a).

299 *Id.* § 7(b); Sabatino, *supra* note 2, at 1243.

300 UHCDA § 2.

301 *Id.* § 1; see Sabatino, *supra* note 2, at 1242.

302 UHCDA § 2(b).

303 *Id.* § 2(b) & cmt.

304 *Id.*

305 *Id.* § 5(a).

306 *Id.* § 5(b).

307 Charles P. Sabatino, *The Evolution of Health Care Advance Planning Law and Policy*, 88 MILBANK Q. 211, 217 (2010).

308 Dunlap, *supra* note 14, at 378 (noting that states enacted penalties against people who coerce a patient into or out of executing a mental health directive); Lester J. Perling, *Health Care Advance Directives: Implications for Florida Mental Health Patients*, 48 U. MIAMI L. REV. 193 (1993).

medication.³⁰⁹ The potential for undue influence may be why almost all states with separate mental health directive statutes have included restrictions on who may serve as a witness.³¹⁰ Several states prohibit members of a principal's family and treatment team members from serving as witnesses.³¹¹ Typically, witnesses must attest to certain observations such as that the principal executed the directive voluntarily.³¹²

State statute approaches underscore the fact that protections against abuse are necessary. The Uniform Act's minimal execution requirements expose mental health patients to risks of undue influence, fraud, and coercion. Because the Uniform Act does not allow for Ulysses arrangements, this section criticizes the Uniform Act's minimal execution requirements for any mental health directive. Ensuring that there are robust protections against abuse is even more critical in the context of Ulysses arrangements. This is because there is a danger that family and providers will coerce the patient into forming a Ulysses arrangement and then use the arrangement as a tool to intimidate the patient to comply with a treatment regimen.³¹³

Elimination of Witnessing Requirements Poses Undue Risks

The Uniform Law Commissioners' elimination of a witness requirement removes an important protection against undue influence, coercion, and fraud.³¹⁴ Witness attestation that the principal showed identification or that the witness knew the principal and had no reason to suspect the principal executed the directive under undue influence or fraud helps ensure execution was voluntary. Moreover, the Commissioners should remove from the potential witness pool people who may have conflicts of interest. Allowing family and treatment team members, who often hold strong opinions about optimum treatments, to witness the directive presents unnecessary risks of coercion and undue influence.³¹⁵ Agents should not serve as witnesses because they have the authority to make all

309 Bruce J. Winick, *Outpatient Commitment: A Therapeutic Jurisprudence Analysis*, 9 PSYCH. PUB. & L. 107 (2003).

310 VHA Report, *supra* note 10, at 5 (stating that concerns over coercion and undue influence caused all of the states with separate statutes except Montana to restrict who may serve as witnesses).

311 *Id.* (listing Arizona, Hawaii, Idaho, Illinois, Kentucky, Michigan, New Mexico, North Carolina, Oklahoma, Oregon, Pennsylvania, South Dakota, Tennessee, Texas, Utah, and Washington as excluding family members and all of the previously listed states in addition to Wyoming as excluding treatment team members).

312 *Id.* (stating most mental health directive statutes require witness attestation except Indiana, Maine, Maryland, Montana, and Washington).

313 *Id.*

314 Stith, *supra* note 2, at 47–48 (“The streamlined procedures appear to sacrifice safeguards for efficiency.”).

315 VHA Report, *supra* note 10, at 5–6.

health care decisions when the principal lacks capacity.³¹⁶ Furthermore, agents should not witness the same instrument that gives them this power.³¹⁷ People affiliated with health care facilities in which the principal receives treatment should not serve as witnesses because they have financial interests in administering care.

Elimination of the Signed Writing Requirement Removes Safeguards

The Uniform Act has been commended for permitting patients to orally designate surrogates and issue treatment instructions because this flexibility is practical and removes obstacles to advance health care planning.³¹⁸ Most people do not create written directives, possibly because they do not like to think about death. When patients do issue instructions, they tend to do so informally.³¹⁹ The typical patient may say, "If I lose capacity, my daughter should make decisions concerning my care."³²⁰ The Uniform Act enforces oral instructions and designations of surrogates for this reason.³²¹

This is why the Uniform Act's elimination of the signed writing requirement is another example of its focus on end-of-life circumstances, not episodic mental illness. For patients with mental illness, the risks posed by enforcing oral instructions and designations of surrogates do not justify the purported benefits. This is because oral instructions are less portable, more susceptible to fraud, and make physicians vulnerable to false accusations. Moreover, requiring a signed writing better ensures the patient has capacity when he forms a directive.

First, people with mental illness need portable instructions and designations of agents. Oral instructions are not as readily portable as a written directive.³²² Patients often receive treatment for acute episodes in emergency rooms or in prison health clinics.³²³ Mental illness patients benefit from portable directives that can be followed wherever they receive treatment.

Second, oral instructions are less reliable and more susceptible to misinterpretation and fraud than written directives. Because mental illnesses are complicated,³²⁴ patient instructions in such cases are often nuanced. If a patient makes off-the-cuff remarks under the stress of an impending crisis, the Uniform Act grants the physician, who has financial interests in administering care and strong opinions about optimal treatments, the authority to record and therefore

316 UHCDA § 1(2).

317 *Id.*

318 See Sabatino, *supra* note 2, at 1244–45.

319 FURROW, *supra* note 100, at 849 (stating that only 10–25% of Americans have documented end-of-life choices or appointed an agent).

320 Sabatino, *supra* note 2, at 1244–45.

321 *Id.*

322 Sabatino, *supra* note 2, at 1243.

323 See *supra* note 29.

324 DSM-V, *supra* note 166.

interpret these inherently unreliable oral remarks.³²⁵ Enforcing oral instructions risks misinterpretation of patient wishes. Moreover, it creates opportunities for health care fraud. Mental health care is particularly susceptible to fraud because: (1) strict patient confidentiality makes abuse hard to discover; (2) the practice of mental health medicine is highly subjective; and (3) mental health patients are often less able to chronicle their treatment than other patients.³²⁶ Therefore, in the mental health sector, enforcing oral instructions recorded by the patient's physician increases opportunities for fraud and abuse.³²⁷

Third, a signed writing requirement protects doctors from fraudulent claims that they administered treatment without informed consent. The written, signed directive documents informed consent to all treatment administered pursuant to its terms. A physician's notes recording a patient's oral remarks may not provide sufficient evidence of informed consent if the patient claims she never consented.

Finally, determining the precise moment an episode causes a person to lose capacity required to issue binding oral instructions is difficult. There is a concern, for example, that a patient with bipolar disorder will utter oral instructions to his doctor when he is hypomanic but technically has capacity. During some episodes of mental illness, patients refuse treatment even though they would have requested treatment if they were not altered by an episode.³²⁸ Moreover, acute episodes may alter a person's judgment in other ways. A bipolar patient when hypomanic may associate with people with whom he would not associate when he was well and ask one of these strangers to be his surrogate. Requiring a written directive better ensures that patients have full capacity when they issue instructions or designate an agent.

The Capacity Presumption and Definition

The Uniform Act defines capacity as an individual's ability to understand the significant benefits, risks, and alternatives to proposed health care and to make and communicate health care decisions.³²⁹ There is a rebuttable presumption of capacity.³³⁰ No clinician determination of capacity is necessary to create a directive.³³¹ The Uniform Act's definition of capacity is similar to the definition of capacity used in many mental health directive statutes.³³² Most states' statutes, like the Uniform Act, have a statutory presumption of capacity to execute any

³²⁵ UHCDA § 7(b).

³²⁶ See Pamela H. Bucy, *Health Care Fraud and the False Claims Act*, ABA CTR. CONTINUING LEGAL EDUC. NAT'L INST., Nov. 19–20, 1998, at *9 (1998), available at N98CFCB ABA-LGLED E-1 (Westlaw); Rosenfeld, *supra* note 114, at 77.

³²⁷ *Id.*

³²⁸ See *supra* note 10 and accompanying text.

³²⁹ UHCDA § 1(3).

³³⁰ *Id.* § 11(b).

³³¹ *Id.*

³³² See, e.g., HAW. REV. STAT. §§ 327G-2, -5 (2013).

directive, including a mental health directive.³³³ However, in Louisiana, an individual wishing to create a mental health directive, but not a generic directive, must obtain a clinician's written attestation that she examined the principal and determined the principal had capacity.³³⁴

Requiring a physician attestation of patient capacity to form any mental health directive³³⁵ stigmatizes people with mental illness and creates an unnecessary administrative obstacle to advance planning. When free from the influence of an episode, many people with mental illness are no less able to make rational treatment choices than people who do not have a mental illness.³³⁶ This is why the Commissioners' decision to use the same definition of capacity for the physical and the mental health contexts makes sense.³³⁷ Moreover, unlike Louisiana, the Commissioners wisely decided to presume that all patients have capacity to form a directive. That presumption is appropriately rebuttable because some patients lack capacity.

However, there are compelling reasons to require a capacity determination at the time a patient forms a Ulysses arrangement. Ulysses arrangements are instruments of self-paternalism.³³⁸ They respect patient autonomy by empowering the patient to direct her health care even during episodes that destroy capacity.³³⁹ It is essential that the patient has capacity when she forms a Ulysses arrangement. Critics worry that because mental illness often negatively affects loved ones, family may be prone to coerce the patient into forming the arrangement.³⁴⁰ A physician's attestation of the principal's capacity at the time of execution is necessary to ensure self-binding treatment is what the patient really wants.

D. Activation

To choose the best activation standard, it is necessary to identify the available options: (1) legal incompetence (used for guardianship proceedings), (2) decision-making capacity (used for informed consent),³⁴¹ (3) dangerousness or severe disability (used for involuntary commitment), or (4) patient-designated activation.³⁴²

333 Winick, *supra* note 95, at 68 n.39; Sheetz, *supra* note 9, at 413.

334 LA. REV. STAT. ANN. § 28:224 (2013); Sheetz, *supra* note 9, at 414.

335 See *supra* note 326 and accompanying text.

336 Sheetz, *supra* note 9, at 405.

337 UHCDA § 1(3).

338 See Dresser, *supra* note 180, at 851.

339 See Willigenburg and Belaere, *supra* note 8, at 395–96.

340 See Winick, *supra* note 95, at 87.

341 See, e.g., UHCDA § 1(3).

342 Sheetz, *supra* note 9, at 414–15; UHCDA § 2.

Courts determine legal incompetence, but physicians determine incapacity.³⁴³ A legal incompetence activation standard does not empower patients to prevent damage caused by mental illness for two reasons. First, a legal incompetence activation standard vests judges with the authority to determine when the directive becomes active even though judges have no specialized training in mental illness or in evaluating a patient's mental state.³⁴⁴ Second, the legal incompetence activation standard obstructs the patient's ability to obtain care. If a court determination is required before a directive becomes active, many patients will not be able to obtain intervention in time. Because physicians, not courts, determine capacity, selecting incapacity as the default activation standard when patients fail to designate one better serves patients.³⁴⁵

An involuntary commitment activation standard is even more problematic for patients attempting to obtain early intervention than an incompetence standard because a person cannot be committed involuntarily unless he is dangerous or gravely disabled a very high threshold.³⁴⁶ This strict standard delays intervention.³⁴⁷ Capacity is an appropriate default standard for a Ulysses arrangement because the arrangement's purpose is to obtain early intervention and avoid commitment. For directives which refuse treatment, early activation ensures that doctors follow patient wishes despite the fact that an episode may obstruct the patient's ability to express refusals.

Admittedly the incapacity activation standard has its drawbacks. First, capacity is often fluid and difficult to determine for patients with mental illness.³⁴⁸ Second, a physician capacity determination takes time. However, the delays are not nearly as long as those caused by court hearings and rulings.

One of the only strengths of the Uniform Act in the mental health context is its decision to allow patients to determine the triggers that allow their directives to take effect. This decision facilitates early intervention even more than an incapacity activation standard. If the patient does not designate a different circumstance, the power of attorney for health care becomes effective when the primary physician determines the patient has lost capacity.³⁴⁹ The Uniform Act commentary uses the following example to illustrate patient designated activation.³⁵⁰ A mother may not want to continue to make her own health care

343 Sheetz, *supra* note 9, at 415; Jessica Wilen Berg et al., *Constructing Competence: Formulating Standards of Legal Competence to Make Medical Decisions*, 48 RUTGERS L. REV. 345, 345-49 (1996).

344 *Washington v. Harper*, 494 U.S. 210, 210, 231-32 (1990) (concluding that a person's interests are better served by allowing the decision to medicate to be made by a medical professional rather than an untrained judge); Sabatino, *supra* note 2, at 1245 (asserting that the Commissioners wanted to keep most health care decisions out of court).

345 See Sheetz, *supra* note 9, at 401.

346 See *supra* Part I.B.

347 *Id.*

348 See *supra* notes 105-106 and accompanying text.

349 UHCDA § 2(c).

350 *Id.* §2(c) cmt.

decisions and may prefer that her daughter make them for her. This mother may specify that her daughter should immediately have power of attorney, even before the mother becomes incapacitated. The mother retains the right to revoke the power of attorney at any time if she does so in writing.

Unlike the Uniform Act, the majority of states with separate mental health directive statutes do not allow directives to become active until the patient has lost capacity.³⁵¹ However, like the Uniform Act, some states empower a person to create a mental health directive which takes effect before loss of capacity.³⁵²

Although the Commissioners adopted the standard that best empowers patients, patient designated activation, the Commissioners should provide more guidance on patient designated activation in the mental health context. Scholars contend activation before loss of capacity is important in the mental health context because early activation enables patients to prevent crisis.³⁵³

The following is an example of a patient-designated activation clause:

My bipolar disorder follows a pattern. Normally, I take my medication and remain stable. However, stress can make me lose sleep, which causes me to become hypomanic. When I am hypomanic, I no longer recognize my need for treatment and stop my medication. While I'm hypomanic, it is possible that my physician may determine I still technically possess capacity. Left untreated, my condition will deteriorate until I become psychotic.

This directive shall become active when my daughter and brother execute a signed affidavit, listing observed symptoms and attesting that they have concluded that I have become hypomanic. I have decided not to require a court determination of my incompetence or a physician's determination of my incapacity to activate this directive because such a requirement would delay treatment.

This patient recognizes that if he chooses incapacity, as determined by his psychiatrist as the activation standard, all of the following will have to take place before treatment. Someone will have to transport the patient to his psychiatrist. He will resist. Someone will have to make an appointment with his psychiatrist. Even if he obtains an appointment, when the patient is hypomanic, he will not

351 VHA Report, *supra* note 10, at 8 (stating that in the 19 states with separate statutes, directives don't become active until the patient loses capacity).

352 ME. REV. STAT. TIT. 18-A, § 5-802 (2013); N.J. STAT. ANN. § 26:2H-108 (WEST 2013); N.M. STAT. ANN. § 24-7B-4 (2013); 20 PA. CONS. STAT. ANN. § 5824 (2013); WASH. REV. CODE ANN. § 71.32.060 (2013); *see also* VHA Report, *supra* note 10, at 8 (asserting that the states previously listed allow for activation before incapacity).

353 Janet Ritchie et al., *Advance Directives in Psychiatry: Resolving Issues of Autonomy and Competence*, 21 INT'L. J.L. & PSYCHIATRY 245 (1998); Gary N. Sales, *The Health Care Proxy for Mental Illness: Can It Work and Should We Want It To?* 21 BULL. AM. ACAD. PSYCHIATRY & L. 161 (1993); VHA Report, *supra* note 10, at 8.

want treatment. He will try to convince his psychiatrist of his capacity. The psychiatrist may not be accustomed to working with directives³⁵⁴ and will be reluctant to hospitalize the patient pursuant to the directive. Instead, the patient has listed two people he trusts and has required them to sign an affidavit attesting to their observations. For him, this strikes the right balance between protection against undue influence and obtaining early intervention. The patient should be free to make this choice.

This activation clause is considered an early activation clause because it activates the directive before a physician has determined he has lost capacity.³⁵⁵ Critics argue early activation is problematic because it creates potential for coercion, which in this instance is brought on by the power vested in family members to activate the directive. However, every patient should be free to create an individualized plan.³⁵⁶ This freedom results in a lack of standardization of directives. Therefore, the patient, his physician, and the trusted family members will need to be sure of what the plan entails. A patient concerned about family having too much control can rely on the presumptive activation standard of a physician determination of incapacity.³⁵⁷ The Uniform Act, which authorizes patient designated activation, clarifies that a patient with capacity may override her directive or the instructions of her agent. This is another protection against coercion.³⁵⁸

E. Expiration

The Commissioners did not provide for automatic expiration of directives. Rather, directives expire under their own terms or when principals revoke them. However, many state mental health directive statutes³⁵⁹ provide for automatic directive expiration after a specified time frame, usually somewhere between two and five years. No state legislature imposes this arbitrary expiration³⁶⁰ on generic directives.

Advocates of automatic expiration assert the following rationale. First, as technology evolves, treatment options change.³⁶¹ A patient's mental illness evolves over the course of the patient's life. Automatic expiration of mental health directives ensures directives continue to reflect patients' treatment instructions over time as their illnesses and treatment options evolve. Second, automatic expiration requires patients to engage in ongoing dialogue with doctors

354 See *supra* note 295 and accompanying text.

355 VHA Report, *supra* note 10, at 8.

356 Sheetz, *supra* note 9, at 403.

357 UHCDA § 2(c)–(d) & cmt.

358 VHA Report, *supra* note 10, at 4.

359 See VHA Report, *supra* note 10, at 10; see, e.g., OHIO REV. CODE ANN. §2135.03 (West 2013); OR. REV. STAT. ANN. § 127.702 (2013); TENN. CODE ANN. § 33-6-1003 (2013); TEX. CIV. PRAC. & REM. CODE ANN. § 137.002(b) (West 2013).

360 VHA Report, *supra* note 10, at 10.

361 *Id.*

and periodically reassess treatment instructions.³⁶² This therapeutic process results in improved decision-making.

Despite these purported benefits, automatic expiration of mental health directives imposes a burden only on patients with mental illness. This burden is unjustified, impractical, and unfair.³⁶³ First, imposing automatic expiration only on mental health directives unjustifiably treats patients with mental illness differently than other patients. The alleged policy reasons supporting automatic expiration of mental health directives apply equally to generic directives. Technology constantly evolves for end-of-life treatment just as it does for mental health treatment. Patients planning for end-of-life care would also benefit from ongoing dialogue with their physicians.³⁶⁴ Just as treatment preferences and goals evolve over time for patients with mental illness, preferences change for the terminally ill.

Second, automatic expiration poses significant administrative burdens only on patients with mental illness who are forced to track the age of directives and re-execute directives every couple of years.³⁶⁵ Few patients execute mental health directives.³⁶⁶ It is likely that even fewer patients would monitor the age of their mental health directives.

Third, requiring automatic expiration only for mental health directives stigmatizes patients with mental illness and undermines parity for mental health care.³⁶⁷ The National Ethics Committee of the Veterans Health Administration indicated that it knew of no evidence supporting the proposition that instructions in a mental health directive “are less stable” than patient instructions in a generic directive.³⁶⁸

However, the automatic expiration administrative burden is necessary *only for* Ulysses arrangements, not revocable mental health directives, for the following reasons. First, automatic expiration helps address concerns about unanticipated consequences due to a patient’s change of heart or failure to foresee all contingencies. Automatic expiration after a few years would require the patient to reaffirm her decision to have a Ulysses arrangement. Reaffirmation would help ensure that the patient continues to want physicians to override her illness-induced refusals of treatment.

Second, automatic expiration only of Ulysses arrangements helps address concerns critics raise that consent provided in a Ulysses arrangement is not valid informed consent because it is not contemporaneous. These critics argue that when doctors treat a patient in accordance with her directive, despite

362 *Id.*

363 *Id.*

364 *See id.*

365 *Id.*

366 *See supra* note 295 and accompanying text.

367 VHA Report, *supra* note 10, at 4.

368 *Id.*

contemporaneous objections, the doctors rely on expired consent.³⁶⁹ Automatic expiration does not guarantee that consent is happening at the moment of treatment, but it does help ensure that the patient has given consent to her directive relatively recently. This is because automatic expiration requires the patient to reaffirm the directive every couple of years.³⁷⁰

F. Advance Consent to Intrusive Treatments, the Role of Proxies, and Patients Who Fail to Plan

This section describes mental health treatments a person might address in a directive. It explores how the Uniform Act addresses advance consent to intrusive treatments, the selection and authority of proxies, and situations when a patient fails to plan. Next, it surveys state approaches to selection of proxies and advance consent to intrusive treatments. Finally, this section evaluates the Uniform Act approach to conclude that it poses undue risks to patients.

Mental Health Treatments

A basic understanding of mental health treatments is necessary to evaluate whether the Uniform Act provides sufficient guidance on advance consent to such treatments. Doctors did not use psychiatric (also known as psychotropic) drugs to treat mental illness until the late 1940s after the discovery that lithium effectively treated bipolar disorder.³⁷¹ Antipsychotic medications are a class of psychiatric medications doctors began using to treat psychosis in the 1950s.³⁷² In the beginning, antipsychotic medications proved effective in limiting psychosis.³⁷³ Censuses in state psychiatric hospitals dropped in the years following the widespread use of antipsychotic drugs.³⁷⁴ It soon became obvious that while antipsychotic medication minimized psychosis, it also potentially caused serious side effects.³⁷⁵ Because of the side effects of various psychiatric medications, courts and legislators consider psychiatric medication to be an intrusive treatment.³⁷⁶ In the 1990s, the United States Food and Drug Administration approved some new antipsychotic drugs for treating patients with

369 See *supra* notes 187–189 and accompanying text.

370 See e.g., OH. REV. CODE ANN. § 2135.03 (2013); TENN. CODE ANN. § 33-6-1003(a) (2013); TEX. CIV. PRAC. & REM. CODE ANN. § 137.002(b) (West 2013).

371 See SLOBOGIN ET AL., *supra* note 47, at 23.

372 Douglas Mossman, *Unbuckling the “Chemical Straitjacket”: The Legal Significance of Recent Advances in the Pharmacological Treatment of Psychosis*, 39 SAN DIEGO L. REV. 1033 (2002).

373 Perlin, *supra* note 54, at 400.

374 *Id.* at 398.

375 *Id.*

376 *Washington v. Harper*, 494 U.S. 210, 221–22 (1990).

psychotic disorders.³⁷⁷ Although these medications are not always effective and do not cure the illness, they are possibly more effective than the older antipsychotic medications.³⁷⁸ The new drugs are “atypical” because they are different than the older antipsychotic medications in that they alleviate psychotic symptoms with fewer side effects.³⁷⁹

Electroconvulsive therapy (ECT), generally considered to be a more invasive treatment than drug therapy,³⁸⁰ directs electric currents to parts of the brain, which induces a series of seizures.³⁸¹ There is not yet a scientific consensus on the explanation for the purported therapeutic benefits of ECT.³⁸² Historically, patient advocates criticized ECT because of its side effects, such as memory loss, dental trauma, bone fractures, and skin burns.³⁸³ Today, improved technology for administering ECT combined and improvements in muscle relaxants have resolved many of the side effects.³⁸⁴ However, certain side effects remain, including memory loss, which can result in permanent memory gaps.³⁸⁵ Although the modern psychiatric community recognizes ECT as an effective and safe treatment for patients who suffer from severe depression and a viable alternative for patients unable to take medication or for whom medication is ineffective, the community is not in unanimous agreement with critics saying that ECT is ineffective and can damage the brain.³⁸⁶

The California legislature defines psychosurgery as including operations referred to as lobotomy, psychiatric surgery, behavioral surgery,³⁸⁷ or any surgery performed to modify or control thoughts, feelings, or behavior, rather than treat a known, diagnosed physical disease of the brain.³⁸⁸ Doctors used prefrontal lobotomy for decades to treat depression, bipolar disorder, obsessive-compulsive disorder, and schizophrenia.³⁸⁹ Today, the medical community considers prefrontal lobotomy to be a discredited, dangerous treatment whose benefits are

377 Mossman, *supra* note 372, at 1039 (The new drugs are clozapine, risperidone, olanzapine, quetiapine, and ziprasidone.).

378 *Id.*

379 *Id.*, at 1039–40.

380 *See infra* note 427 and accompanying text.

381 Mike E. Jorgensen, *Is Today the Day We Free Electroconvulsive Therapy?*, 12 QUINNPIAC HEALTH L.J. 1, 3–4 (2008).

382 Helia Garrido Hull, *Electroconvulsive Therapy: Baby Boomers May Be in for the Shock of Their Lives*, 47 U. LOUISVILLE L. REV. 241, 251 (2008).

383 Jorgensen, *supra* note 381, at 10.

384 SLOBOGIN ET AL., *supra* note 47, at 27.

385 Hull, *supra* note 382, at 254–56; *In re Estate of Austwick*, 656 N.E.2d 779, 781 (Ill. App. Ct. 1995) (listing fractures, memory loss, confusion, delirium, and in rare instances, death as side effects).

386 Hull, *supra* note 382, at 251, 259.

387 *See* CAL. CODE WELF. & INST. § 5325 (West 2013).

388 *Id.*

389 Henry T. Greeley, *Neuroscience and Criminal Justice: Not Responsibility but Treatment*, 56 U. KAN. L. REV. 1103, 1111–12 (2008).

outweighed by the significant risks of permanent brain damage.³⁹⁰ Modern psychosurgery techniques are referred to as stereotactic procedures and involve creating small lesions in different areas of the brain.³⁹¹ Generally, even these modern procedures are rarely used and considered highly intrusive.³⁹² When performed, the procedures are typically restricted to hospitalized patients with very serious mental disorders only after less intrusive therapies have failed.³⁹³

The Uniform Act Approach

Health care decisions of the guardian, agent, or surrogate are effective without judicial approval.³⁹⁴ If neither the patient nor the court has designated a proxy, or the proxy is unavailable, any available family member may act as surrogate under a priority system starting with the spouse.³⁹⁵ The surrogate must promptly inform the other family members of her assumption of authority.³⁹⁶ Unless the person is related to the patient, an owner, operator, or employee of a residential long-term health care institution at which the patient receives care may not act as a surrogate or agent.³⁹⁷ The patient may disqualify a person from acting as her surrogate by a signed writing or by informing her doctor.³⁹⁸ The Uniform Act imposes no other safeguards against the nomination of proxies who might depart from patient wishes.³⁹⁹

If members of a class of surrogates who have equal priority (i.e. siblings) disagree about a treatment decision, majority rule applies.⁴⁰⁰ If these surrogates are still evenly divided, the decision-making process stops⁴⁰¹ and a court-appointed guardian makes the decision.⁴⁰² Agents and surrogates must make decisions in accordance with the patient's instructions or known wishes.⁴⁰³ If there are no instructions, agents and surrogates shall make decisions pursuant to the patient's best interests while considering the patient's values.⁴⁰⁴

A patient can consent to mental health treatment in a directive by issuing instructions and/or designating a proxy, thereby avoiding the time consuming

390 *Id.*

391 *Id.*

392 SLOBOGIN ET AL., *supra* note 47, at 32.

393 *Id.*; 1 KAPLAN & SADOCK'S COMPREHENSIVE TEXTBOOK OF PSYCHIATRY 1914 (Benjamin J. Sadock et al., eds., 9th ed. 2009).

394 UHCDA §§ 2(f), 5(g), 6(c); *see also id.* § 6(b).

395 *Id.* § 5(b).

396 *Id.* § 5(d).

397 *Id.* § 5(i).

398 *Id.* § 5(h).

399 *See Stith, supra* note 2, at 58.

400 UHCDA § 5(e) & cmt.

401 *Id.*

402 *Id.* §§ 5(e) & cmt., 14; *see Sabatino, supra* note 2, at 1249.

403 UHCDA §§ 2(e), 5(f).

404 *Id.*

process of resorting to a court-appointed proxy.⁴⁰⁵ Whether a patient can grant a surrogate or agent the authority to consent to inpatient mental health treatment has been the subject of debate and was the impetus for the 1999 amendment to the Uniform Act.⁴⁰⁶ As amended, the Uniform Act now prohibits agents and surrogates from consenting to the patient's inpatient mental health treatment unless the written directive expressly provides such authority.⁴⁰⁷

Despite the requirement for express written authorization for an agent or surrogate to consent to inpatient mental health treatment, the Uniform Act does not seem to require express written authorization for an agent or surrogate to consent to outpatient ECT or psychotropic medication.⁴⁰⁸ This is because the Uniform Act's broad grant of authority to agents and surrogates includes the authority to make all health care decisions for the patient.⁴⁰⁹ Health care decisions include selection and discharge of doctors, approval and disapproval of tests, surgeries, medications, and orders to resuscitate, and directions to provide, withhold, or withdraw artificial nutrition, hydration, and other care.⁴¹⁰ This definition appears broad enough to encompass ECT, psychotropic medication, and even psychosurgery. However, the Uniform Act and its commentary do not explicitly mention ECT, psychotropic medication, or psychosurgery. Rather, the enumerated examples of health care decisions focus on end-of-life decisions.

Under the Uniform Act, clinicians and institutions have the duty to comply with a patient's instructions as well as a proxy's decisions.⁴¹¹ Decisions of agents, surrogates, and guardians obligate the clinician or institution to the same extent as the patient's instructions.⁴¹² Override provisions delineate limited instances in which the physician need not follow the directive.⁴¹³ First, clinicians and institutions may refuse to implement instructions or a proxy's decisions for reasons of conscience.⁴¹⁴ Second, clinicians and institutions may refuse to implement instructions or a proxy's decisions requiring medically ineffective care or treatment contrary to accepted standards.⁴¹⁵ When the clinician or institution refuses to treat in accordance with the instruction or proxy decision,

405 *Id.* § 2.

406 *Cohen v. Bolduc*, 760 N.E.2d 714, 613, 614 nn.13 & 14 (Mass. 2002); *see also* Winick, *supra* note 95, at 81–86.

407 UHCDA § 13(e) (clarifying that UHCDA does not address whether a guardian has authority to consent to the principal's inpatient mental health treatment but leaving that matter to state guardianship law); *see also id.* § 13(f) (stating that UHCDA does not affect mental health treatment of involuntarily committed people because state law addresses this matter).

408 *Id.* §§ 1(2), 1(6), 1(17), 2, 5.

409 *Id.* §§ 1(6), 2(b), 5(a).

410 *Id.* § 1(6).

411 *Id.* § 7(d).

412 *Id.*

413 *Id.* § 7(e)–(f); *see* VHA Report, *supra* note 10, at 6.

414 UHCDA § 7(e).

415 *Id.* § 7(f).

they must notify the patient or her authorized representative.⁴¹⁶ Moreover, they must make reasonable efforts to assist in transferring the patient to another facility willing to comply with the directive and provide continuing care until transfer.⁴¹⁷

State Approaches

Most states provide broader protection against health care fraud than the Uniform Act does, by removing people affiliated with any facility treating the patient.⁴¹⁸ Some states that have implemented the Uniform Act have amended the Uniform Act's automatic surrogate selection priority system to safeguard against selecting a surrogate who might depart from patient values.⁴¹⁹ For example, Delaware disqualifies a spouse when there has been a complaint of domestic abuse.⁴²⁰ Hawaii refused to enact the priority list and selects surrogates based on consensus of interested parties.⁴²¹

Whether a principal can convey authority to an agent to consent to the principal's admission in a mental health facility depends on the principal's state of residence. Some states authorize patients to create mental health directives but do not allow patients to empower an agent to consent to inpatient mental health treatment.⁴²² In North Dakota, patients are allowed to convey an agent authority to consent to voluntary commitment of the patient for up to 45 days but are not allowed to consent to a commitment for any greater length of time.⁴²³ Other states follow the Uniform Act approach by allowing an agent to consent to the principal's inpatient treatment only with express authority in a written directive.⁴²⁴ Finally, some states allow an agent to consent to the principal's inpatient mental health treatment even without express commitment authority as long as the grant of authority is sufficiently broad.⁴²⁵

Whether a patient may use a directive to consent to intrusive treatments also depends on the principal's state of residence. Many states prohibit patients from consenting to or conveying authority to an agent to consent to psychosurgery in a

⁴¹⁶ *Id.* § 7.

⁴¹⁷ *Id.*

⁴¹⁸ See, e.g., ALASKA STAT. § 13.52.010(c) (2013); CAL. PROB. CODE § 4659(a)–(b) (2013).

⁴¹⁹ Stith, *supra* note 2, at 58.

⁴²⁰ DEL. CODE ANN. TIT. 16, § 2507(b)(2)(f) (2013).

⁴²¹ Stith, *supra* note 2, at 58.

⁴²² See, e.g., TEX. HEALTH & SAFETY CODE ANN. § 166.152(f)(1) (West 2013); WIS. STAT. ANN. § 155.20(2) (West Supp. 2013); *Cohen v. Bolduc*, 760 N.E.2d 714, 714, 718 n.15 (Mass. 2002).

⁴²³ N.D. CENT. CODE § 23-06-5-03 (2013); *Cohen*, 760 N.E.2d at 718 n.15.

⁴²⁴ See, e.g., ARIZ. REV. STAT. § 36-3283(f) (West Supp. 2013); FLA. STAT. ANN. § 765.113(1) (West 2013); HAW. REV. STAT. § 327E-13(e) (2013); MISS. CODE ANN. § 41-41-227(5) (Lexis Nexis 2013); *Cohen*, 760 N.E.2d at 719 n.17.

⁴²⁵ See, e.g., *Cohen*, 760 N.E.2d at 723.

directive.⁴²⁶ Several states do not empower a principal to convey authority, even expressly, to an agent to consent to the principal's ECT; a court order is required.⁴²⁷ Kentucky empowers patients to issue binding refusals of treatments but, arguably, does not empower patients to issue binding consents.⁴²⁸ This is because Kentucky explicitly authorizes patients to use directives to refuse specific medications or ECT but only to state preferences for medications or emergency interventions.⁴²⁹ This language suggests doctors are bound to adhere to the patient's refusals and must consider patient medication preferences but are not required to administer those medications.⁴³⁰ Finally, some states empower a principal to use a directive to consent to and convey authority to an agent to consent to intrusive treatments, including ECT⁴³¹ and psychotropic medication.⁴³²

Analysis

Insufficient Protection under the Uniform Act's Proxy Limitations

The Uniform Act limitation of the potential agent and surrogate pool does not protect patients with mental illness, particularly against health care fraud. There is evidence that health care fraud is more pervasive in the mental health sector.⁴³³ Therefore, patients with mental illness are more vulnerable than other patients to receiving treatment they do not need or to which they have not consented. A person who has a financial incentive in administering treatment

426 See, e.g., CAL. PROB. CODE § 4652 (West 2013); OR. REV. STAT. § 127.540 (2013); TEX. HEALTH & SAFETY CODE ANN. § 166.152 (West 2013); WASH. REV. STAT. § 11.92.043 (2013) (prohibiting a guardian from consenting to surgery solely for the purpose of psychosurgery).

427 See Jorgensen, *supra* note 381, at 1 (stating that many states require proxies to obtain prior court authorization before consenting to ECT on behalf of an incapacitated ward); see, e.g., CAL. PROB. CODE § 4652 (West 2013); DC STAT. § 7-1231.07(e) (2013); N.H. REV. STAT. § 464-A:25 (2013); OR. REV. STAT. § 127.540 (2013); TEX. HEALTH & SAFETY CODE ANN. § 166.152(f) (2013).

428 See KY. REV. STAT. § 202A.422 (West 2013); Sheetz, *supra* note 9, at 425.

429 KY. REV. STAT. § 202A.422 (West 2013).

430 *Id.*

431 ECT is more regulated than psychotropic medication. See Jorgensen, *supra* note 381, at app. A (providing a table of state statutes concerning ECT); see, e.g., MICH. COMP. LAW § 330.1717 (2013) (prohibiting administration of ECT without consent from the patient, the guardian, or the agent if the directive grants the agent authority to consent to ECT); WASH. REV. STAT. §§ 71.32.260, 71.32.160 (2013) (allowing the principal to indicate whether she consents and authorizes her agent to consent to administration of ECT).

432 See, e.g., IND. CODE § 16-36-1.7-3 (2013) (authorizing a patient to specify in a directive psychotropic medication, electroconvulsive therapy, and inpatient treatment); MINN. STAT. § 253B.03(6)(d) (2013); N.C. GEN. STAT. ANN. § 122C-73 (West 2013) (allowing use of a directive to grant or withhold authority for psychotropic medication, electroconvulsive therapy, and inpatient mental health treatment); WASH. REV. STAT. §§ 71.32.100, 71.32.050, 71.32.260 (2013) (authorizing a patient to make a declaration consenting to or refusing intrusive mental health treatments and to convey authority to a proxy to make decisions about intrusive mental health treatments).

433 See *supra* note 326 and accompanying text.

should not serve as the patient's agent, unless she is a family member. The Commissioners recognized that patients in nursing homes are particularly vulnerable and therefore prohibited people affiliated with the patient's nursing home from being the patient's surrogate or agent.⁴³⁴ However, this should not be the only limitation on the potential surrogate pool because most patients receive mental health care outside of long-term residential health care institutions. Unless the person is related to the principal, no owner, operator, employee, or agent of *any* facility where the principal receives care should act as an agent or surrogate.

Too Much Authority to Automatically Selected Surrogates

Although the Uniform Act's definition of health care decisions appears broad enough to encompass ECT, psychotropic medication, or even psychosurgery, the Uniform Act and its commentary never specifically address any of these intrusive treatments.⁴³⁵ This vacuum of guidance combined with the Uniform Act's broad grant of authority to surrogates the principal never chose⁴³⁶ poses undue risks of coercion and undue influence.

The following story illustrates how the Uniform Act's lack of guidance on mental health treatments combined with broad authority to automatically selected surrogates poses undue risks. In the past, Ms. Jones alleged that Mr. Jones abused her. She voluntarily admits herself in the psychiatric ward because she is severely depressed. After admission her psychiatrist determines that she has lost capacity. She has no guardian, agent, or directive. In the past, Ms. Jones has taken medication to treat her mental illness. She has never expressed any opinion about ECT. Mr. Jones notifies her psychiatrist that he will serve as her surrogate. The psychiatrist explains treatment options to Mr. Jones which include outpatient ECT. Mr. Jones selects ECT to treat his wife. Her siblings and parents disagree. Although they have never discussed ECT with Ms. Jones, they believe she would not want to receive it. They also do not trust Mr. Jones. Nonetheless, Mr. Jones authorizes ECT.

Unless administering ECT violates the doctor's conscience, is an ineffective treatment, or contrary to accepted standards (triggering the Uniform Act override provision), the doctor must comply with the patient's decision and administer ECT.⁴³⁷ Most likely, none of these narrow exemptions apply to ECT, which, despite its side effects, is recognized as an effective treatment. Even if the doctor refuses to administer ECT under the override provision, the doctor must make reasonable efforts to transfer Ms. Jones to a facility willing to administer ECT.⁴³⁸ The Uniform Act requirement for Ms. Jones's express written authorization to

434 Sabatino, *supra* note 2, at 1243–44.

435 UHCDA § 1(6).

436 *Id.* § 5(a).

437 *Id.* § 7(d)–(f).

438 *Id.* § 7(g).

enable a proxy to consent to her inpatient mental health treatment does not limit Mr. Jones's power to consent to his wife's outpatient ECT.⁴³⁹

Although Mr. Jones must have express authority to consent to his wife's inpatient mental health treatment, the Commissioners failed to impose such a requirement on outpatient psychotropic medication, ECT, or psychosurgery.⁴⁴⁰ Because Ms. Jones left no instructions, Mr. Jones is supposed to make her health care decisions based on his estimation of what is in her best interests.⁴⁴¹ Based on their history, it is quite possible that he would not make decisions in her best interests. Invoking one of the few Uniform Act safeguards, one of Ms. Jones's family members may petition the court for an injunction to stop administration of ECT.⁴⁴² However, if no family member cares enough to do so or if the family member does not prevail, Mr. Jones's decision controls the course of treatment. The history of domestic violence does not limit his power to make health care decisions for Ms. Jones because the Commissioners neglected to remove from the surrogate pool family members who might depart from patient values.⁴⁴³ Neither Ms. Jones nor a court evaluated whether Mr. Jones could be trusted to make decisions in line with his wife's values. Most likely, Ms. Jones would not have chosen Mr. Jones.

It is appropriate for the Uniform Act to create a decision-making framework for the vast majority of the population who fails to plan for end-of-life care.⁴⁴⁴ Turning to family makes sense.⁴⁴⁵ However, one size does not fit all. In the mental health context, arbitrarily selecting the spouse to be surrogate because he is first in line undermines patient autonomy. The Uniform Act's grant of unchecked authority to a single family member whom the patient never chose ignores the realities of mental illness which often devastates familial relationships, especially marriages.⁴⁴⁶ Even in the absence of domestic abuse, many patients would not want their spouses, acting alone without court approval, to have the power to authorize ECT, psychotropic medication, or psychosurgery. The Uniform Act's failure to remove from the surrogate pool individuals who might depart from patient values undermines patient autonomy and risks coercion and undue influence.

Look To Patient's Grant of Authority

When a surrogate is automatically selected, the patient has no input.

439 *Id.* § 13(e) & cmt.

440 *Id.* § 13(e).

441 *Id.* § 5(f).

442 *Id.* § 14.

443 Stith, *supra* note 2, at 57–58.

444 See Sabatino, *supra* note 2, at 1248–49.

445 FURROW, *supra* note 100, at 849.

446 Davoli, *supra* note 10, at 1045 (recommending early intervention before a mental illness erodes a patient's support system).

Providing broad authority to an agent the patient selected is less problematic. The patient maintains influence over her care.⁴⁴⁷ When the patient has capacity, she can communicate her preferences to her agent. She can select an agent she trusts to make decisions consistent with her values. However, the Uniform Act does not authorize an agent to consent to the principal's inpatient mental health treatment based on the principal's grant of unlimited health care decision-making authority to the agent.⁴⁴⁸ The Act prohibits an agent from consenting to the principal's admission to a mental health care institution unless the principal's written directive expressly provides such authority.⁴⁴⁹ *Cohen v. Bolduc* underscores the reasons why this arbitrary limitation undermines patient autonomy.⁴⁵⁰

In *Cohen v. Bolduc*, the Massachusetts Supreme Court analyzed whether Massachusetts' general health care proxy statute authorized an agent to commit a principal to a mental health facility when the principal did not oppose.⁴⁵¹ The principal's health care proxy stated:

My Health Care Agent is granted full power and authority to consent to any and all medical treatment which I may need in the event that I am unable to consent . . . including without limitation authority to consent to medical care, hospitalization, nursing home admission, or whatever else may in my Health Care Agent's sole judgment be in my best interest I further state . . . that there are no limitations imposed upon my Health Care Agent's authority.⁴⁵²

The proxy was activated when Bolduc's psychiatrist decided Bolduc lacked capacity.⁴⁵³ Bolduc's psychiatrist admitted Bolduc into a mental health facility under Massachusetts's emergency psychiatric hospitalization procedures. Bolduc's agent then converted Bolduc's admission status to conditional voluntary, a status which imposed no temporal limits on Bolduc's hospitalization. Had her agent not done so, Massachusetts law would have required the hospital to file a petition to retain Bolduc involuntarily which would have required proof that Bolduc met strict involuntary commitment criteria. Later, Bolduc revoked her proxy and demanded discharge from the hospital.⁴⁵⁴

The Massachusetts statute did not address whether the principal's grant of unlimited decision-making authority conveyed to the agent the authority to consent to the principal's inpatient mental health treatment.⁴⁵⁵ However, the

447 See Winick, *supra* note 95, at 82–85.

448 UHCDA § 13(e).

449 *Id.*

450 *Cohen v. Bolduc*, 760 N.E.2d 714, 720–24 (Mass. 2002).

451 *Id.* at 715.

452 *Id.*

453 *Id.* at 716.

454 *Id.* at 716–17.

455 *Id.* at 718.

proxy statute granted an agent the authority to make any health care decisions for the principal and defined “health care” broadly to include treatment of “mental conditions.”⁴⁵⁶ *Cohen* concluded that the statutory language suggested that agents had the authority to commit the principal to a mental health facility.⁴⁵⁷

Cohen also considered the policy implications and wisely determined that prohibiting an agent from committing her principal frustrated the purpose of the proxy statute to support patient autonomy.⁴⁵⁸ Under the statute, the agent’s decisions had the same effect as the principal’s decisions. *Cohen* correctly stated that restricting the range of advance planning choices unduly limited the principal’s ability to control her own care. The *Cohen* court departed from the Uniform Act by not requiring express authority in a written directive to empower the agent to commit the principal.⁴⁵⁹ The principal’s grant of unlimited authority to make health care decisions was deemed sufficient.

For a directive to be an effective tool, a patient must be able to use the directive to consent to mental health treatments and empower an agent to do the same. The Uniform Act’s requirement for express authority in a written directive for an agent to consent to inpatient mental health treatment⁴⁶⁰ undermines patient autonomy for the following reasons.

First, arbitrary limitations on an agent’s ability to consent to the principal’s treatment will result in principals not receiving care they need and to which they consented. The Uniform Act requires the principal to appreciate that she must use “magic” words conveying authority to an agent to consent to her admission in a mental health institution.⁴⁶¹ It is illogical that Bolduc’s written grant of unlimited authority to make health care decisions to her daughter would not include the right to consent to admission in a mental health facility.⁴⁶² However, had the Massachusetts Supreme Court applied the Uniform Act, this would have been the illogical result.⁴⁶³ Patients who grant unlimited authority to agents have the right to expect that doctors and agents will look to the patient’s own words to determine the scope of the agent’s authority.

Second, limiting a patient’s right to consent, in advance, to inpatient and pharmacological mental health treatment imposes a unique burden on patients with mental illness who have been historically stigmatized. States provide patients the authority to refuse life-sustaining treatment, either through instructions or through agents.⁴⁶⁴ The U.S. Supreme Court has recognized that the right of patient autonomy can outweigh the significant state interest in

456 *Id.* at 720.

457 *Id.*

458 *Id.* at 721.

459 *Id.* at 715.

460 UHCDA § 13(e) & cmt.

461 *Id.*

462 *Cohen*, 760 N.E.2d at 715.

463 UHCDA § 13(e) & cmt.

464 See FURROW, *supra* note 100, at 829–31.

preservation of life.⁴⁶⁵ No state interest in preserving life is implicated when a patient grants an agent authority to consent to psychotropic medication and inpatient mental health treatment. On the other hand, there is a tremendous patient interest in securing treatment to prevent a crisis.

However, ECT is a unique, invasive treatment.⁴⁶⁶ The Uniform Act fails to mention ECT and can be construed to authorize a surrogate the patient never selected to consent to the patient's outpatient ECT.⁴⁶⁷ Moreover, under the Uniform Act, a patient who conveys broad decision-making authority but never mentions ECT arguably conveys authority to an agent to authorize the patient's ECT.⁴⁶⁸ ECT is more invasive and controversial than pharmacological therapy.⁴⁶⁹ It is for this reason that many states prohibit an agent from consenting to ECT without a court order.⁴⁷⁰ Because ECT is an effective treatment for many patients, patients should be able to consent to and convey authority to an agent to consent to the patient's ECT.⁴⁷¹ Considering the invasive nature of ECT, advance consent to ECT should have to be expressed.

No Proxy Consent to Psychosurgery

The Uniform Act grants the authority to make all health care decisions for the principal when the principal lacks capacity to patients' agents, surrogates, and guardians.⁴⁷² This authority is unchecked because no court approval is necessary.⁴⁷³ The Uniform Act's enumerated examples of health care decisions focus on end-of-life decisions and do not mention psychosurgery.⁴⁷⁴ On its face, the broad definition of health care decision appears to include psychosurgery.⁴⁷⁵ Granting a proxy this authority is unwise, which may be why many states prohibit a proxy from consenting to the principal's psychosurgery.⁴⁷⁶ Instead, doctors should use other, less intrusive treatments to restore the patient's capacity. At this point, when the patient has regained full capacity, the patient can decide whether to consent to more invasive treatments such as psychosurgery.

465 See *supra* notes 151–153 and accompanying text.

466 See *supra* notes 380–386 and accompanying text.

467 UHCDA § 1(6).

468 *Id.* §§ 1(6), 2(b).

469 See *supra* notes 380–386 and accompanying text.

470 See *supra* note 427 and accompanying text.

471 See *supra* notes 384 and 386.

472 UHCDA § 1(6).

473 *Id.* §§ 2(f), 5(g), 6(c).

474 *Id.* § 1(6)(iii) (listing orders concerning artificial nutrition and hydration).

475 *Id.* § 1(6)(ii) (including approval and disapproval of surgical procedures).

476 See *supra* note 426 and accompanying text.

G. Directive Templates

The Uniform Act provides a model statutory form which allows principals to designate agents and provide instructions.⁴⁷⁷ Principals may check boxes to indicate whether they want to prolong life as long as possible. One section of the form allows the principal to indicate whether she wants artificial nutrition and hydration withheld and whether she wants to donate her organs.

The Uniform Act template fails to address common mental health issues, but more importantly, showing this form at the wrong time to already vulnerable patients could be confusing and unsettling to the patient. One commentator noted that some hospitals are reluctant to provide patients with templates that ask for instructions about harvesting organs because it gives the wrong impression at the wrong time.⁴⁷⁸ This is especially true for the patient who has recently recovered from a mental illness episode and wants to plan for a future episode. Having to address end-of-life issues distracts the patient from the mental illness issues he faces and could deter him from creating a directive.

There are real benefits to offering a template. First, the process of developing a template involves consulting with experts. Stakeholders have an opportunity to resolve potential issues. This process makes it easier for courts to uphold the directive when problems arise. Moreover, providing a preapproved form arguably encourages patients to execute directives. It is simpler for a person to check boxes than it is for the person to craft an individualized health care plan.

However, some scholars argue that the drawbacks of a model form outweigh the benefits.⁴⁷⁹ One scholar observed that “statutory forms tend to become fixed realities with a life of their own that is resistant to change.”⁴⁸⁰ Patients with bipolar disorder, depression, schizophrenia, drug addiction, or various other mental illnesses⁴⁸¹ may elect to create mental health directives. Even patients diagnosed with one mental illness, such as bipolar disorder, will present in different ways and experience different levels of severity of the illness.⁴⁸² A single template cannot address the unique needs of the varied patient population. Psychiatrists should encourage patients to create their own directives tailored to their individual needs. There is a danger that a statutory form will become the standard from which patients are afraid to deviate.⁴⁸³

477 UHCDA § 4.

478 Sabatino, *supra* note 2, at 1248.

479 See *supra* notes 480–483.

480 Sabatino, *supra* note 2, at 1248.

481 DSM-V, *supra* note 166, at 87–122 (schizophrenia), 123–88 (mood disorders), 481–590 (substance-related disorders).

482 *Id.* at 123–54.

483 See Sabatino, *supra* note 2, at 1248.

III. SOLUTION: ISSUE A MODEL STATUTE

This Part recommends the Uniform Law Commissioners adopt a model mental health directive statute, because the Uniform Act, with its focus on end-of-life, fails people with mental illness. The Uniform Law Commission is the appropriate organization to issue a model statute because its mission is to promote uniform state laws in areas where uniformity is desirable and practical. In the area of advance planning for mental health treatment, patients benefit from uniformity.

In half of the states that have not enacted mental health directive statutes, patients must rely on generic directive statutes that fail to address important mental health issues. Even in the states with specialized statutes, provisions vary widely and often do not meet the needs of patients with mental illness.⁴⁸⁴ A person's ability to control her illness should not depend on the state in which she resides. Moreover, patients frequently receive care for acute episodes away from their hometowns. Uniform direction from the Commissioners helps ensure that a patient's directive is valid wherever she receives care.

The model provisions in Appendix A improve on the Uniform Act and the Washington statute. The model provisions fall into the following categories:

- (A) Provisions making mental health directives, particularly Ulysses arrangements, effective intervention tools;
- (B) Provisions ensuring mental health directives, particularly Ulysses arrangements, are created free from undue influence, coercion, or fraud;
- (C) Safeguards ensuring mental health directives, particularly Ulysses arrangements, are properly implemented and not abused; and
- (D) Provisions removing obstacles to advance planning.

A. Provisions Making Directives Effective Intervention Tools

Several of the model provisions are designed to empower patients to use their directives to control their illnesses. The model activation provision provides guidance the Uniform Act fails to give on how patient designated activation works in the mental health context. Many state mental health directive statutes do not allow patients to determine the standard by which their directives become active. Postponing activation until the point at which a physician determines the patient has lost capacity delays care. The model language allows patients to designate when their directives become active and facilitates early intervention.

The Uniform Act imposes the arbitrary requirement for express authority in a written directive for an agent to authorize the principal's inpatient mental health treatment. This burden could deprive people of care they need and to which they consented. Under the model provisions, the plain language of the principal's written grant of authority determines the scope of the agent's authority to consent

⁴⁸⁴ See *supra* Part II.

to the principal's inpatient mental health treatment and psychotropic medication. This meets the expectations of a principal who grants unlimited authority to an agent to make her health care decisions.

The Uniform Act fails to address ECT. Some states prohibit an agent from consenting to the principal's ECT. Because ECT is recognized as an effective treatment, the model provisions empower patients to use directives to consent to ECT and to grant authority to agents to do the same.

The model language enables a patient to form a Ulysses arrangement. First, the model legislative findings recognize that issues implicated in end-of-life care differ from issues confronting patients with mental illness and recognizes the value of Ulysses arrangements. The Uniform Act and even the Washington statute prevent people from forming truly self-binding arrangements for care.⁴⁸⁵ The model provision empowers patients to receive three weeks of inpatient mental health care pursuant to their irrevocable directives despite illness-induced refusals.

The model language better enables patients to control their illnesses than the Washington statute does in two ways. First, the model language eliminates the Washington statutory requirement for discharge of an incapacitated patient who demands discharge, even if discharge contravenes her irrevocable directive.⁴⁸⁶ Second, unlike the Washington statute,⁴⁸⁷ the model language creates a rebuttable presumption of incapacity when a patient's irrevocable directive consents to treatment that the patient then refuses under the influence of an episode. This rebuttable presumption recognizes that episodes often cause patients to refuse treatment. Concerned about liability for unlawfully involuntarily treating a patient, doctors will likely adjudge patients as having capacity when episodes cause patients to refuse treatment requested in their directives.⁴⁸⁸ This caution harms the patient because it prevents intervention. The rebuttable presumption encourages doctors to follow the directive, despite the patient's illness-induced objections.

B. Provisions Ensuring Directives are Created Free from Undue Influence, Coercion, or Fraud

Provisions Applicable to all Mental Health Directives

The Uniform Act's elimination of the requirement for a signed, witnessed writing⁴⁸⁹ removes protections for patients who are vulnerable to undue influence, coercion, and fraud. The model execution provision contains

485 WASH. REV. CODE § 71.32.140(6)(b) (2013).

486 *Id.*

487 *See supra* Section II.B.

488 *Id.*

489 UHCDA §§ 2(a), 5(a).

safeguards the Uniform Act eliminates such as the requirement of a signed writing witnessed by two disinterested people who attest that the principal presented identification and did not appear coerced.

Provisions Applicable Only to Ulysses Arrangements

Recognizing the sensitivities of administering treatment despite illness-induced refusals, the model provisions create safeguards to ensure Ulysses arrangements are formed voluntarily by patients with capacity. First, the model provisions give every patient the right to choose whether their directives will be revocable when they lack capacity. No patient should have as her only advance planning option an irrevocable directive. Unlike most state mental health directive statutes, which require capacity to revoke,⁴⁹⁰ the model language allows patients to revoke their directives even when they lack capacity unless they designate otherwise.

Patients who enter Ulysses arrangements must obtain a clinician attestation of patient capacity. Moreover, Ulysses arrangements automatically expire two years after formation unless the patient reaffirms. Automatic expiration helps address concerns about the risks of unintended consequences due to a patient's change of heart or failure to foresee all contingencies. Physicians and patients must engage in ongoing dialogue to ensure the directive remains consistent with the patient's wishes. Automatic expiration also addresses concerns that consent provided in a Ulysses arrangement is not contemporaneous informed consent. Admittedly consent provided through the arrangement is not truly contemporaneous, but automatic expiration ensures that consent is relatively recent.

C. Safeguards Ensuring Directives are Properly Implemented and Not Abused

Safeguards Applicable to all Mental Health Directives

Unlike the Uniform Act, which only excludes from the agent/surrogate pool people affiliated with long-term residential facilities,⁴⁹¹ the model language focuses on mental health. Many people receive mental health treatment outside nursing homes. To protect against fraud, which is more common in the mental health sector,⁴⁹² the model language removes from the potential agent pool people affiliated with *any* facility treating the patient.

The model activation provision protects against coercion by clarifying that a directive does not prevail over contemporaneous preferences of a principal with capacity. The model language also prohibits a directive from authorizing

490 See VHA Report, *supra* note 10, at 9.

491 UHCDA §§ 2(b), 5(i).

492 See Bucy, *supra* note 326.

psychosurgery, a controversial and rarely used treatment.⁴⁹³ The Uniform Act's lack of guidance on intrusive mental health treatments combined with broad authority granted to surrogates the principal never selected⁴⁹⁴ poses undue risks of coercion. In the model statute, the Commissioners should eliminate provisions for automatic surrogate selection. People whom neither the principal nor a court selected should not have the power to authorize intrusive mental health treatments.

Because ECT is more controversial and invasive than pharmacological therapy, the model language protects patients by forbidding an agent from consenting to the principal's ECT unless the principal's directive expressly grants such authority.

Safeguards Applicable to Ulysses Arrangements

The model language sets forth safeguards that address concerns that Ulysses arrangements violate due process because they enable clinicians to forcibly hospitalize and treat a patient even when the patient does not meet commitment criteria. First, doctors must heed treatment refusals from patients with capacity. Second, the model language requires express consent before administering psychotropic medication in contravention of illness-induced objections. Moreover, only licensed psychiatrists may administer the medication and only if two psychiatrists recommend in writing the specific medication. Third, the provisions limit self-binding hospitalization to three weeks. Fourth, although patients may consent in a revocable directive to ECT and authorize an agent to do so, the provisions do not allow Ulysses arrangements for ECT. This is because administering ECT, an invasive treatment, in contravention of contemporaneous objections, creates undue opportunities for coercion. Instead, the patient can use the Ulysses arrangement to obtain intervention, through other means, such as drug therapy. Then, when the patient regains capacity, she can consent to ECT if she so chooses. Fifth, the model language allows patients to seek injunctive relief. All of these safeguards help address concerns that Ulysses arrangements create opportunities for undue influence, abuse, and coercion from doctors and family desperate to conquer the patient's illness.

D. Provisions Removing Obstacles to Advance Planning

Some states create obstacles to directive formation that further stigmatize mental illness, such as requiring a capacity determination to create any mental health directive.⁴⁹⁵ The model language creates no unnecessary obstacles to

493 See *supra* Section II.F.

494 UHCDA § 5(a).

495 See *supra* note 334 and accompanying text.

directive formation. Patients need only obtain a capacity determination if they form a Ulysses arrangement.

The Uniform Act contains a template that addresses end-of-life issues and could be confusing and upsetting to otherwise healthy patients with mental illness. This inappropriate template could deter patients from forming a mental health directive. Instead of a template,⁴⁹⁶ the Commissioners should provide a range of samples of directives tailored to the needs of patients with different mental health planning needs. Samples would assist the patient create a directive tailored to her own needs. Multiple samples would avoid creating the impression that deviation from a mandatory form could render the directive unenforceable.

CONCLUSION

Ulysses arrangements enable people to obtain intervention when an acute mental illness episode prevents them from recognizing they need treatment. The Uniform Act purports to be a comprehensive model advance directive statute, which addresses all types of advance health care planning,⁴⁹⁷ but it fails to meet the needs of people with mental illness, most notably by failing to empower patients to form Ulysses arrangements. Washington's approach is touted as being at the forefront of patient empowerment.⁴⁹⁸ However, even Washington prohibits Ulysses arrangements by requiring discharge of an incapacitated patient who demands discharge even when releasing the patient contravenes her irrevocable directive.⁴⁹⁹

The Uniform Law Commissioners should issue a model statute that empowers patients to enter Ulysses arrangements, removes roadblocks to directive formation, creates parity for mental health care, and prevents fraud, coercion, and undue influence. The recommended provisions accomplish these goals. Unlike Washington, the model language does not require doctors to heed a patient's illness-induced discharge demands, which are in contravention of her directive.⁵⁰⁰ Further, the model language creates a rebuttable presumption of incapacity in the event that a patient's irrevocable directive consents to treatment that the patient then refuses under the influence of an episode.⁵⁰¹ This presumption facilitates treatment because it recognizes doctors will be reluctant to treat a patient in the face of illness-induced refusals. If the Commissioners adopt this model statute and states follow suit, people with mental illness will have more power to control their own treatment.

496 UHCDA § 4.

497 *See* UHCDA § 3.

498 *See* WASH. REV. CODE § 71.32.140 (2013); Sheetz, *supra* note 9, at 401, 433.

499 WASH. REV. CODE § 71.32.140(6)(b) (2013).

500 *Id.* § 71.32.140(6)(b).

501 *Id.* § 71.32.140(2)(a).

* * *

APPENDIX A: MODEL STATUTORY PROVISIONS GOVERNING MENTAL HEALTH DIRECTIVES

*Legislative Findings*⁵⁰²

(1) *Issues implicated in advance planning for end-of-life care are distinct from issues implicated in advance planning for mental health care.*

(2) *An individual with capacity has the right to control decisions relating to her mental health care.*

(3) *Mental illness is often episodic. Periods of incapacity obstruct the individual's ability to give informed consent and impede the individual's access to mental health care.*

(4) *Facilitating advance planning helps: (a) prevent unnecessary involuntary commitment and incarceration, (b) improve patient safety and health, and (c) improve care and enable patients to exercise control over their treatment.*

(5) *An acute episode can induce an individual to refuse treatment when the individual would consent to treatment if the individual's judgment were unimpaired. Empowering people to create self-binding mental health advance directives ("directives") to overcome their illness-induced treatment refusals protects patient safety, autonomy, and health.*

(6) *Individuals with mental illness have the same rights to plan in advance for treatment as individuals planning for end-of-life care. A directive can only accomplish the goals listed above if a patient may use a directive to:*

(a) *Set forth instructions for mental health care, including consent to inpatient mental health treatment, psychotropic medication, or electroconvulsive therapy;*

(b) *Dictate whether the directive is revocable during periods of incapacity and consent to treatment despite illness-induced refusals;*

(c) *Choose the standard by which the directive becomes active; and*

(d) *Designate an agent to make health care decisions for the patient.*

*Execution of Directives*⁵⁰³

A directive shall:

(1) *Be in writing;*

(2) *Be dated and signed by the principal or the principal's designated representative if the principal is unable to sign;*

(3) *State whether the principal wishes to be able to revoke the directive at*

502 See *id.* § 71.32.010 (2013). The Washington statute inspired these model findings, which emphasize Ulysses arrangements even more than Washington does.

503 These provisions were inspired by WASH. REV. CODE. §§ 71.32.050, 71.32.060, 71.32.090 (2013). Unlike Washington, however, this provision requires a mental health professional attestation of principal capacity to form a Ulysses arrangement and allows principals who fail to address revocation to freely revoke. See *id.* § 71.32.070.

any time or whether the directive remains irrevocable during periods of incapacity. Failure to clarify whether the directive is revocable does not render it unenforceable. If the directive fails to state whether it is revocable, the principal may revoke it at any time.

(4) Contain a principal affirmation that the principal is aware of the nature of the document signed and signed the directive freely and voluntarily;⁵⁰⁴

(5) Be witnessed in writing by at least two adults. No witness may be:

(a) A member of the principal's treatment team;

(b) Related to the principal by blood, adoption, or marriage;

(c) Be in a romantic or dating relationship with the principal;

(d) The agent of the principal or a person designated to make health care decisions for the principal; or

(e) The owner, operator, employee, or relative of an owner or operator of a treatment facility in which the principal is a patient.

(6) Witnesses shall attest:

(a) They were present when the principal signed the directive;

(b) The principal did not appear incapacitated or under undue influence or duress when the principal signed the directive; and

(c) The principal presented identification or the witness personally knows the principal.

(7) Contain a written, signed attestation from a mental health professional that the principal had capacity at the time of directive execution, but only if the principal makes the directive irrevocable. If the principal is free to revoke the directive at any time, no mental health professional attestation of principal capacity is required.

(8) Be valid upon execution.

Activation of Directives

(1) Activation is the point at which the directive is used as the basis of decision making and dictates treatment of the principal.⁵⁰⁵

(2) Unless the principal otherwise designates in the directive, a directive becomes active when the principal loses capacity.

(3) The principal may designate an activation standard other than incapacity by describing the circumstances under which the directive becomes active.

(4) Despite activation, a directive does not prevail over contemporaneous preferences expressed by a principal who has capacity.

Role of Agents⁵⁰⁶

⁵⁰⁴ See *Hargrave v. Vermont*, 340 F.3d 27, 32 n.2 (2d Cir. 2003) (citing VT. STAT. ANN. tit. 14, § 3456 (2002) (which inspired this provision)).

⁵⁰⁵ VHA Report, *supra* note 10, at 8.

⁵⁰⁶ See WASH. REV. CODE § 71.32.100, which inspired this provision. However, this model provision addresses an agent's role in a Ulysses arrangement.

(1) *In a directive, a principal may appoint an agent to make all health care decisions for the principal, including decisions to consent on behalf of the principal to electroconvulsive therapy, inpatient mental health treatment, and psychotropic medication.*

(2) *Express authorization to the agent to consent to the principal's inpatient mental health treatment and/or psychotropic medication is not required to convey authority to an agent to consent to such treatments. Rather, the agent may consent to such treatments for the principal if the principal's written grant of authority is sufficiently broad to encompass these decisions. However, an agent only has the authority to consent to electroconvulsive therapy for the principal if the principal expressly granted authority to consent to the principal's electroconvulsive therapy.*

(3) *An agent's decisions for the principal must be in good faith and consistent with the principal's instructions expressed in the principal's directive. If the directive fails to address an issue, the agent shall make decisions in accordance with the principal's instructions or preferences otherwise known to the agent. If the agent does not know the principal's instructions or preferences, the agent shall make decisions in the best interests of the principal.*

(4) *If the principal grants the agent authority to make decisions for the principal in circumstances in which the principal still has capacity, the principal's decisions when the principal has capacity override the agent's decisions.*

(5) *Except as otherwise prohibited by law, an agent has the same right as the principal to receive, review, and authorize the use and disclosure of the principal's health care information as is necessary for the agent to carry out the agent's duties for the principal.*

(6) *Health care decisions an agent makes for a principal are effective without judicial approval.*

(7) *When an incapacitated principal refuses inpatient mental health treatment and/or psychotropic medication, the principal's agent only has the authority to consent to such treatments for the principal if the principal's irrevocable directive expressly authorizes the agent to consent to the applicable treatment.*

(8) *A principal may not designate as her agent an owner, operator, or employee of a facility at which the principal is receiving care or a relative of such owner or operator unless the designated person is related to the principal by blood, marriage, or adoption.*

Permissible Scope of Directives

In directives, principals may issue instructions or appoint agents to make decisions concerning all aspects of their mental-health treatment, except as limited by subsection (4) below, including:

(1) Consent to or refusal of specific types of mental health treatments, including psychotropic medication, electroconvulsive therapy, and inpatient

mental health treatment; Consents to electroconvulsive therapy must be express;

(2) Preferences concerning treatment facilities and care providers;

(3) Nomination of a guardian for the court to consider if guardianship proceedings commence; but

(4) Principals may not consent to or authorize agents to consent to psychosurgery in a directive.

*Revocation of Directives; Procedures for Implementing Self-Binding Arrangements, and Automatic Expiration only of Self-Binding Arrangements*⁵⁰⁷

(1) Except self-binding arrangements as described in (5) below, directives remain in effect until they expire under their own terms or are revoked by the principal. Self-binding arrangements automatically expire two years after they are executed unless the principal reaffirms the arrangement. In the event the principal is incapacitated at the end of the two-year time frame, the self-binding arrangement remains in effect until the principal regains capacity and determines whether to reaffirm the arrangement.

(2) A principal may freely revoke a directive even if she is incapacitated unless the principal makes her directive irrevocable during periods of incapacity. To be irrevocable, the directive shall:

(a) State that the directive remains irrevocable during periods of incapacity; and

(b) Contain an attestation from a mental health professional that the principal had capacity at the time of executing the directive.

(3) A principal with capacity or a principal without capacity who did not make her directive irrevocable during periods of incapacity may revoke a directive by:

(a) A written statement revoking the directive;

(b) A subsequent directive that revokes the original directive. If the subsequent directive does not revoke the original directive in its entirety, only inconsistent provisions in the original directive are revoked; or

(c) Physical destruction of the directive with the intent that it be revoked.

(4) When a principal with capacity consents to treatment that is different than the treatment requested in her directive or refuses treatment that the principal requested in her directive, this consent or refusal does not revoke the entire directive but is a waiver of the inconsistent provision.

(5) A principal has a right to form a self-binding arrangement for care. Self-binding arrangements allow the principal to obtain treatment in the event that an acute episode renders the principal incapacitated and induces the principal to refuse treatment. To provide advance consent to inpatient treatment despite the principal's illness-induced refusals, in her directive, a principal shall:

(a) Make her directive irrevocable pursuant to subsection (2) above; and

(b) Consent to admission in an inpatient treatment facility.

⁵⁰⁷ Also inspired by *id.*

(c) If the principal wants administration of psychotropic medication despite the principal's illness-induced refusals of medication, the principal shall expressly consent to psychotropic medication in the irrevocable directive.

(6) If the principal forms a self-binding arrangement for treatment but then refuses admission despite the directive's instructions to admit, the facility shall respond as follows:

(a) The facility shall, as soon as practicable, obtain the informed consent of the principal's agent, if any is designated.

(b) Two mental health professionals shall within 24 hours of the principal's arrival at the facility evaluate the principal to determine whether the principal has capacity and to document in the principal's medical record a summary of findings, evaluations, and recommendations.

(c) The principal's statements in her directive requesting inpatient treatment upon activation of the directive, combined with activation of the directive, and contemporaneous refusals of treatment requested in the directive create a rebuttable presumption that the principal lacks capacity.

(d) If the evaluating mental health professionals determine the principal lacks capacity, the principal shall be admitted into the treatment facility pursuant to the principal's directive. The treating mental health professional shall document in the principal's medical records all treatment administered. After 21 days from the date of admission, if the principal has not regained capacity or has regained capacity but refuses to consent to remain for additional treatment, the facility shall release the principal during daylight hours unless the principal is detained pursuant to involuntary commitment standards.

(7) If a principal who has been determined to lack capacity continues to refuse inpatient treatment, the principal may immediately seek injunctive relief for release from the facility.

(8) If a principal with an irrevocable directive consenting to inpatient treatment refuses psychotropic medication through words or actions, only a licensed psychiatrist may administer psychotropic medication, and only if:

(a) The principal expressly consented to psychotropic medication in the principal's irrevocable directive;

(b) The agent, if one was designated, consented to psychotropic medication; and

(c) Two licensed psychiatrists recommend in writing treatment with the specific psychotropic medication.

Refereeing the Public Health

Hosea H. Harvey*

ABSTRACT:

Between January 2009 and October 2013, 49 states and the District of Columbia passed laws focusing on mitigating the consequences of traumatic brain injuries (TBIs) in organized youth sports. Using historical, contextual, and empirical methods, this Article describes the content, goals, and structure of youth sports TBI laws, while hypothesizing about their underlying legislative logic and long-term public health consequences. The Article's empirical evidence suggests two key findings: first, that a dominant interest group, the National Football League, helped to define the problem and its associated solutions for the vast majority of states, thus curving the legislative story arc in favor of its policy prescriptions; second, that existing youth sports TBI laws are focused on secondary, not primary, prevention, and may thus shift attention away from more comprehensive solutions. Finally, the Article explains why such state laws will likely fail to substantially resolve the larger untackled problem—significantly reducing the overall rate and number of TBIs in youth sports. After explaining why existing state youth sports TBI laws fail to accomplish this broader goal, the Article queries whether alternative policy or public health measures might offer more robust solutions.

* Assistant Professor of Law and Political Science, Temple University, James E. Beasley School of Law. Ph.D. (Stanford University), J.D. (Stanford Law School). The author would like to thank The Robert Wood Johnson Foundation and the Center for Health Law, Policy, and Practice at Temple University for research support. The author thanks Temple colleagues and anonymous reviewers for providing helpful comments on an earlier version of this draft and specially thanks Rebecca Cole for her invaluable feedback and attention to detail. The author thanks Shilpa Kadoo, Stephanie Kammer, and Nikos Kaplanov for valuable research assistance. This article benefits from data first reported in Hosea Harvey, *Reducing Traumatic Brain Injuries in Youth Sports: Youth Sports Traumatic Brain Injury State Laws, January 2009-December 2012*, 103 AM. J. PUB. HEALTH 1249 (July 2013).

TABLE OF CONTENTS

INTRODUCTION	69
I. THE EPIDEMIOLOGY OF TBIs AND YOUTH CONCUSSIONS	72
A. TBI DIAGNOSIS AND RECOGNITION: AGREEMENT AND DIVERGENCE ..	72
B. A VAST BUT UNCERTAIN SCOPE	73
C. CONCUSSION ASSESSMENT	75
D. TREATMENT: CONSENSUS AND DIVERGENCE ABOUT SHORT-TERM CONSEQUENCES.....	76
E. IDENTIFICATION AND TREATMENT OF MULTIPLE TBIs	78
F. AMBIGUOUS EVIDENCE LEADS TO AMBIGUOUS POLICY LOGIC.....	79
II. SCIENTIFIC EVIDENCE AND YOUTH SPORTS TBI LAWMAKING	81
III. RUSHING THE FIELD: YOUTH SPORTS TBI LAWMAKING (2009-2012)....	85
A. THE “GALVANIZING INCIDENT” MODEL OF POLICYMAKING.....	85
B. EVALUATING YOUTH SPORTS TBI LAWS—DATA AND METHODOLOGY	88
C. A UNIFORM POLICY SOLUTION?	88
D. THE ROLE OF POLICY EXPERIMENTATION (PURPOSEFUL AND ACCIDENTAL)	95
E. SUBSTANTIVE POLICY INNOVATION: EVIDENCE, THEORY, AND PRACTICE	96
F. EVIDENCE, INFLUENCE, AND RUSHING: COMMON CAUSAL MECHANISMS IN YOUTH SPORTS TBI LAW POLICYMAKING	98
IV. FAILURES OF EXISTING POLICY AND POTENTIAL POLICY IMPROVEMENTS	104
A. MANDATORY AGGREGATE REPORTING AND FEEDBACK REGIMES	105
B. MANDATORY INDIVIDUAL TRACKING METRICS.....	106
C. REQUIRED BASELINES	106
D. DIRECT INTERVENTION IN PARTICULAR SPORTS	107
E. FIXING THE REPORTING DISINCENTIVE.....	110
F. LEARNING FROM PROFESSIONAL SPORTS AND INDEPENDENT	

ASSOCIATIONS	111
G. EVALUATING POLICY ALTERNATIVES	112
CONCLUSION	113
APPENDICES.....	116
APPENDIX 1: EVIDENCE OF INFLUENCE—NEWS.....	116
APPENDIX 2: EVIDENCE OF INFLUENCE—LEGISLATIVE HISTORY	117
APPENDIX 3: NFL STATEMENTS REGARDING ITS OWN LEGISLATIVE INVOLVEMENT	119

INTRODUCTION

In September 2010, Boston University's Center for the Study of Traumatic Encephalopathy ("the Center") found mild states of degenerative brain disease caused by repetitive concussions and/or sub-concussive brain injuries in a former football player. That the Center found such brain-trauma indicators after an athlete's death was not uncommon; its researchers have been at the forefront of studying professional football players for years.² What *was* highly uncommon in this instance was that the Center found the indicators in a recently deceased *college* football player who had tragically committed suicide while still in college.³ This discovery marked the first time that a college player's brain tissue had been systematically tested for the effects of repeated concussions, and it led to renewed efforts to bring concussion education and other reform efforts to youth sports.

In conjunction with another youth football player-related tragedy a year earlier,⁴ efforts to understand the short- and long-term effects of concussions or traumatic brain injuries (TBIs) broadened to include awareness of the effects of multiple traumatic brain injuries (MTBIs) on youth athletes. Because the scope of public awareness was magnified by the tragic experiences of these two young players, as well as publicized injuries in professional sports, a coalition of parents, educators, athletes, and professional groups used these events to marshal various constituencies, including legislatures, to address the problem.⁵ Despite the large number of interest groups galvanized by these events, a single interest group, the National Football League (NFL), soon emerged as the market leader for policy prescriptions addressing youth sports TBIs. With a myriad of statutory

2 The Center is considered the leading research organization focused on developing a comprehensive understanding of the long-term effects of concussions on athletes. Dozens of current and former NFL players are on its brain donation registry, and the NFL promotes this collaboration. *See, e.g., Our Research*, CTR. FOR THE STUDY OF TRAUMATIC ENCEPHALOPATHY, BOS. UNIV., <http://www-test.bu.edu/cste/our-research/> (last visited Dec. 2, 2013).

3 *See* Alan Schwarz, *Suicide Reveals Signs of a Disease Seen in N.F.L.*, N.Y. TIMES, Sept. 13, 2010, <http://www.nytimes.com/2010/09/14/sports/14football.html>.

4 The case involves a Washington state middle-school football player named Zackery Lystedt, discussed below in Part III.A.

5 *See, e.g.,* Ben McGrath, *Does Football Have a Future?*, NEW YORKER, Jan. 31, 2011. Broadly speaking, the high rate of youth concussions and public discussions thereof often focuses on youth football, which does account for a high portion of such injuries. In contrast, except where noted, this Article considers all youth sports where athletes are at risk for TBIs or MTBIs in its examination of the epidemiology and TBI-focused legal regimes. This scope includes all organized youth sports, particularly those with risks of head impact or injury. *See generally* Andrew Gardner et al., *Chronic Traumatic Encephalopathy in Sport: A Systematic Review*, BRIT. J. SPORTS MED. (forthcoming) (manuscript at 1–6) (2013), available at <http://bjsm.bmj.com/content/early/2013/06/25/bjsports-2013-092646.full.pdf> (describing some new understandings of how sports TBIs occur outside of direct football head hits).

and regulatory options available, state laws emerged as the quickest, if not the most effective, solution.

Public health law scholars have engaged in similar policymaking debates, rightly focused on the use of evidence and data when proposing new interventions, particularly those harnessing the force of law. These evidence-based interventions are not unique to public health, and have been widely utilized in other areas.⁶ Further, scholars agree that the ideal environment for implementing laws of the type discussed here would rely on accumulation of existing research, synthesis of that research into a set of discrete outcomes, translation of those outcomes into a policy framework, and an evaluation of that framework's ability to influence and modify the behavior or outcome it seeks to change.⁷ But with respect to youth sports TBI laws, this approach was not consistently followed. Nonetheless, lawmakers acted at a fairly quick pace.

Between January 2009 and October 2013, 49 states and the District of Columbia passed laws designed to minimize the consequences of TBIs in youth sports.⁸ Despite this impressive and sudden legislative response to a long-standing problem, the core logic and evidence supporting such laws has remained a mystery. This uncertainty about the link between science and policy should have led lawmakers to question the efficacy and efficiency of their efforts during this period. What do these laws purport to do? What are the appropriate behavioral and health-science logic models underpinning these legislative mandates? Will these laws have their purported desired effect—a reduction in instances and effects of youth TBIs? And if the laws do eventually have such an effect, how will policymakers, public-health officials, and other advocates know?

This Article is an attempt to answer some of these core questions in the absence of conclusive scientific evidence. By October 2013, although every state but Mississippi had passed some form of youth sports TBI law, policymakers, health advocates, and outside interest groups still had little information as to the expected effectiveness of these laws on reducing the overall *number* of youth TBIs and reducing the long-term *effects* of youth TBIs.⁹ As this Article explains, despite the recent proliferation of state laws, there is no short-term evidence (or

6 See, e.g., Ross C. Brownson et al., *Understanding Evidence-Based Public Health Policy*, 99 AM. J. PUB. HEALTH 1576, 1576–83 (2009).

7 See, e.g., Scott Burris et al., *Making the Case for Laws that Improve Health: A Framework for Public Health Law Research*, 88 MILBANK Q. 169 (2010).

8 For a discussion of laws and data-gathering methods, see Sections III.A and III.B below.

9 Scientific uncertainty notwithstanding, the sudden and swift implementation of the laws and their corresponding lack of evaluative metrics means that no systematic studies were undertaken to test their effectiveness prior to implementation. Therefore, this Article evaluates their theoretical goals, attempts to determine whether those goals can be met, and explores what might be done to further align the goals with existing scientific evidence. The Article focuses primarily on middle- and high-school sports, since almost all existing state laws focus on that domain to the exclusion of others (recreational sports and college sports in particular).

even a metric or framework) to evaluate whether such laws will accomplish their stated goals. Further, because many of these laws follow a familiar blueprint endorsed by the NFL and leading brain injury advocacy groups, there is little substantive policy experimentation between states and so it will be difficult to causally assess differences in state-level outcomes over time. This Article argues that since lawmakers were starting from scratch, they could have aimed for a more rigorous, engaged, and long-term TBI reduction solution, despite some scientific uncertainty about the epidemiology of youth TBIs.

This Article is also a study of a hyperkinetic form of state-level policymaking, and all available evidence suggests that the pace and consistency of such lawmaking can primarily be attributed to the role of outside interest groups, most notably the NFL. Evidence gathered here points toward the NFL's significant role in the promotion of a select group of public health law and policy prescriptions with respect to youth sports TBIs. As a result, it is hard to determine *ex ante* whether laws produced in such an influence-driven environment were designed to actually mitigate youth TBIs or to further the goals of various interest groups, which may not view TBI reduction as their primary objective. Ideally, each state could serve as a laboratory where one centralized, authorized, and motivated state-level regulator would take charge of developing a broad-based public health law and policy solution to combating TBIs in youth populations.¹⁰ No state has created such a centralized regime, and the role of interest groups—the NFL in particular—may have influenced that policy choice. The NFL's intervention may have led states to act, which is likely better than silence. At the same time, the NFL's role may also have caused states to accept an initial slate of recommendations as a final policy prescription, while ignoring nuanced issues of TBI causation and policy alternatives to the NFL's approach. It is not merely that states legislated in a scientific vacuum, but that alternative methodologies or interventions were not considered against the backdrop of the Lystedt Law and the NFL's support.

This Article proceeds as follows. First, the epidemiology of TBIs is analyzed with particular focus on youth populations—referencing adult and college populations when appropriate.¹¹ Next, the Article turns to public health law

10 It is common in public health lawmaking frameworks to analyze the effects of state laws in certain public health areas (tobacco, seat belts, alcohol, etc.) and then to frame the divergence of such laws as an opportunity to test the effectiveness of public health law innovations in the fifty state lawmaking laboratories. For the purposes of this Article, however, the utility of this framework is minimal because of the uniformity of laws across states.

11 The Article does not specifically focus on the robust literature involving collegiate or professional anti-concussion initiatives except as limited comparative frameworks for understanding the general epidemiology of TBIs. Instead, the Article's focus is on the nexus between the science of understanding *youth* TBIs, the legislative responses to that scientific understanding, and the connection between the epidemiology of *youth* TBIs and laws that could, if

intervention efforts at the state level and analyzes youth sports TBI laws for key similarities and differences. The objective in this section is to highlight legal regimes that effectively incorporate the best scientific knowledge—and those that do not. After asking whether these laws are indeed effective in their purported key goal—reducing the incidence and consequences of TBIs in youth populations—the Article analyzes the role of the NFL in the policymaking process. Concluding that the NFL’s role in shaping the legislative story arc was formidable, the Article then turns to analyzing the areas that state youth sports TBI laws chose to ignore, such as evaluative metrics and other direct outcome measures. After examining the failures of the state youth sports TBI laws, the Article turns to addressing a broader question: whether law is the proper forum to address this public health problem. Concluding that existing youth sports TBI laws are indeed necessary and acknowledging that these initial responses did more good than harm, the Article closes by focusing on ways in which new legislative initiatives can improve upon earlier interventions.

I. THE EPIDEMIOLOGY OF TBIs AND YOUTH CONCUSSIONS

A. *TBI Diagnosis and Recognition: Agreement and Divergence*

For purposes of this Article, a concussion or TBI is defined as (a) an injury to the head arising from blunt trauma, an acceleration force, or a deceleration force, (b) which then disrupts the normal functioning of the brain, (c) which causes the individual to exhibit one of many indicators (whether observed or self-reported), and (d) where any such indicator is attributable to the injury.¹² More generally speaking, a concussion is a mild brain injury, caused by trauma to the head, which results in a temporary disruption of normal brain functions.¹³

Since a concussion typically involves some direct or indirect force to the head, observers usually recognize concussions in youth athletes following a perceived head impact or injury. Yet, there is no agreed-upon TBI diagnostic metric, and no uniform national TBI reporting protocols.¹⁴ As a result of this data

properly aligned, minimize their frequency and long-term impact.

12 Collectively, those concussion indicators include the trauma or force followed by (a) transient confusion, disorientation, or impaired consciousness; (b) dysfunction of memory; (c) loss of consciousness; or (d) signs of other neurological or neuropsychological dysfunction including seizures, irritability, lethargy, fuzzy vision, headache, or vomiting.

13 Even this simple definition is not uniformly accepted. Others suggest that body jolts and body blows that do not result in a literal “blow to the head” can cause concussions, because any major body blow can jostle the brain. While this is true, the focus here is on those concussions caused by direct head impacts.

14 Christopher Randolph et al., *Concussion Symptom Inventory: An Empirically Derived Scale for Monitoring Resolution of Symptoms Following Sport-Related Concussion*, 24 ARCHIVES CLINICAL NEUROPSYCHOLOGY 219, 229 (2009).

void, first-line responders to potential TBI events have settled on some commonly accepted—but scientifically imprecise—basic TBI detection methods. At the outset, the traditional method for first recognizing a potential TBI during a youth athletic practice or competition is through visual observation of an impact followed by visible manifestations of various TBI consequences or symptoms. Other players may notice the visual signs of a TBI, such as vomiting, loss of consciousness, or confusion—even when the symptoms are not understood directly by the youth athlete himself. Therefore, one key line of defense in recognizing TBIs is fellow athletes' awareness of typical symptoms.

One step removed, coaches and other team leaders' awareness is often a critical element in TBI diagnosis. To ensure that athletes receive the assistance they require following a TBI, the coach, athletic trainer, and/or other supervisory official should be trained to recognize the various signs and symptoms of a concussion. For example, if a coach does not know that an athlete exhibiting retrograde amnesia or confusion may have a TBI (and may be at risk for another), then the coach may not remove the player from the game, thereby risking further injury. Therefore, a coach's simple visual awareness and alertness on the field, coupled with a fundamental understanding of TBIs and their symptoms, can be critical in diagnosis.

Slightly more removed from initial direct observation, parents are another necessary line of defense. When parents are at competitions watching their child, they can serve as useful monitors to notice when their child or another young athlete is experiencing TBI symptoms. Parents can also be a counterweight in situations where the interests of coaches and players might not be fully aligned. If an injury goes unnoticed on the field, parents are best equipped to notice subtle changes in a young athlete's behavior after the competition and assess whether their son or daughter has experienced a TBI or requires medical attention. Therefore, parents' ability to serve as another safeguard against an injured player's further injury is predicated upon the parents' knowledge of concussive symptoms. Without this knowledge and understanding of TBI symptoms and the risk of further injury, many parents will not understand the symptoms their child may be experiencing; this may cause a potentially dangerous condition to go undiagnosed and untreated.

B. A Vast but Uncertain Scope

No one disagrees that youth are at risk of developing TBIs when they play organized sports. Instead, the discussion about youth and TBIs is characterized by scientific uncertainty about causality, the scope of youth concussion rates, effects of multiple impacts, treatment, evaluation, and assessment of long-term

outcomes.¹⁵ It is difficult to be specific about each of these elements because there is no comprehensive reporting mechanism at the state or national level to identify all instances of youth TBIs during a given year or the percentage of those caused by youth sports.¹⁶ Therefore, it becomes necessary to rely on estimates, anecdotes, and/or incomplete reporting systems to establish baseline population data. When state laws attempt to minimize the number and consequences of TBIs in youth populations, all but a few do so without relying on a state-level evaluative metric or data-driven framework. It is thus difficult to imagine exactly how the success of these laws will be evaluated. Nonetheless, we do know that a problem of some magnitude exists, even if we are unsure of exactly how significant it is.

That said, experts generally believe that the prevalence of sports-related concussions among young people is significantly higher than reported.¹⁷ As identifying metrics and reporting mechanisms have improved, reports of concussions in youth sports have skyrocketed. But because reporting, definitions, treatment standards, and testing have also all improved during the past fifteen years, it is nearly impossible to be certain whether the data trends show an absolute rise or simply reflect more accurate assessment of incidences in youth populations.¹⁸ Most experts agree that children and teenagers are at a greater risk of concussion than adult populations and that concussions in younger people take longer to heal because their brains are still developing. Some estimates suggest that up to 10% of all high-school athletes in contact sports suffer a concussion during each season.¹⁹ Youth sports and bicycle accidents account for the majority of concussion cases among 5- to 14-year-olds.²⁰ And somewhat recent estimates of youth sports concussions suggest that there are between 1.6 and 3 million (or

15 This Article does not purport to review the entire field of literature regarding youth and TBIs. For a brief sampling of work in the field, see Melissa L. McCarthy et al., *Health-Related Quality of Life During the First Year After Traumatic Brain Injury*, 160 ARCHIVES PEDIATRIC & ADOLESCENT MED. 252, 260 nn.12–26 (2006).

16 Although reporting systems have improved over the last few years, the problem for our purposes is tracking longitudinal data about the effects of TBIs and MTBIs over time and across varied and stable populations. Because reporting mechanisms have historically been weak to non-existent, such long-term analysis is not yet possible. See Bryan Jennett, *Epidemiology of Head Injury*, 60 J. NEUROLOGY, NEUROSURGERY & PSYCHIATRY 362, 364 (1996) (noting that “data on admissions for head injury are not routinely collected on a national basis in the United States”).

17 Steven T. DeKosky et al., *Traumatic Brain Injury—Football, Warfare, and Long-Term Effects*, 363 NEW ENG. J. MED. 1293, 1295 (2010).

18 Mike Mitka, *Reports of Concussions From Youth Sports Rise Along With Awareness of the Problem*, 304 JAMA 1775, 1775–76 (2010).

19 See *Concussion Facts*, SPORTS CONCUSSION INST., <http://www.concussiontreatment.com/concussionfacts.html> (last visited Dec. 2, 2013).

20 Allan H. Ropper & Kenneth C. Gorson, *Concussion*, 356 NEW ENG. J. MED. 166 (2007). Bicycling, in addition to being regulated under other laws, is not considered a “youth sport” for the purposes of this Article and shall not be discussed in depth.

more) annually in the United States.²¹

However, because specific data are lacking at both the national and state level, experts acknowledge a high degree of uncertainty notwithstanding their general agreement about the significance of the problem. For example, studies provide estimates ranging from a 200 to 300% increase in reported instances within the last decade.²² Some provide generalized estimates across all sports while others provide estimates specific to certain types of sports.²³ Part of the uncertainty in these estimates is based on the fact that the only uniform national reporting regime, the National Electronic Injury Surveillance System All Injury Program (NEISS-AIP), relies on a highly limited number of reporting hospitals. But the statistics, whether they over- or underestimate, indicate that the problem exists in sufficient magnitude to warrant attention.²⁴

C. Concussion Assessment

Even though most assessments of whether an individual has suffered a TBI are not robustly scientific, this does not mean that more scientific measurements and assessments are impossible. Best practices in the identification of TBIs involve pre-competition baseline measurements of an athlete's overall brain function. Companies, such as ImPACT²⁵ and CogState Sport,²⁶ and healthcare institutions such as the Mayo Clinic²⁷ offer computerized baseline testing. Once baseline information is recorded, the results can be interpreted by comparing them with an injured athlete's post-injury responses to determine if he or she demonstrates cognitive symptoms that would suggest a concussion.²⁸ Generally, developers advertise these tests as low-cost assessment tools that can be administered by a healthcare professional²⁹ or on-site at an athletic facility

21 Jean A. Langlois et al., *The Epidemiology and Impact of Traumatic Brain Injury: A Brief Overview*, 21 J. HEAD TRAUMA REHAB. 375, 375–78 (2006).

22 Ropper & Gorson, *supra* note 20.

23 Luke M. Gessel et al., *Concussions Among United States High School and Collegiate Athletes*, 42 J. ATHLETIC TRAINING 495, 497 (2007) (providing both overall and sport-specific estimates).

24 There are many more studies spanning decades of research. This Article seeks to summarize broad research themes without purporting to be the definitive scientific review of all extant TBI literature.

25 See ImPACT, <http://www.impacttest.com> (last visited Dec. 2, 2013).

26 See CogState, <http://www.cogstate.com> (last visited Dec. 2, 2013).

27 See Susan Shepherd, *Mayo Clinic Offers Baseline Concussion Testing to Arizona High School Athletes*, MAYO CLINIC NEWS, June 23, 2011, <http://newsblog.mayoclinic.org/2011/06/23/mayo-clinic-offers-baseline-concussion-testing-to-arizona-high-school-athletes>.

28 *Id.*

29 See *Frequently Asked Questions*, ImPACT, <http://impacttest.com/about/?Frequently-Asked-Questions-7> (last visited Dec. 2, 2013).

through the use of a computer or tablet device.³⁰ These computerized tools provide coaches and staff with another diagnostic metric that can be used to prevent youth athletes with TBIs from prematurely returning to play and risking further injury. For example, the Mayo Clinic's test can be taken from any computer with Internet access, takes 8 to 15 minutes to complete, and allows the athlete or parent to share the results with their chosen healthcare providers. These tools are not meant to be used in isolation, and their results may require health provider intervention. The Mayo Clinic recommends use of its test in conjunction with a comprehensive medical evaluation when an athlete is suspected of having a concussion.³¹ In those cases where baseline tests reveal abnormalities or where other indicators point to TBI complications, the athlete would then be directed to receive medical treatment and preventive measures. Notwithstanding the provider of the test or its ease of use, for a variety of reasons—cost likely among them—few states require baseline testing for any of their student athletes.

D. Treatment: Consensus and Divergence about Short-Term Consequences

Once an athlete is suspected of having a TBI, removal from competition is obviously the first immediate form of treatment and prevention of further injury. However, while there is some initial consensus about immediate first steps, there is substantial disagreement about actions that should follow.

Generally speaking, initial treatment of a TBI involves assessing the severity of the injury and monitoring the athlete's condition. Depending on the severity of the injury, many athletes will require rest until their concussive symptoms have subsided. Recommendations found in the *National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion* provide that a coach, athletic trainer, or physician should monitor an athlete with a concussion at "5-minute intervals from the time of the injury until the athlete's condition completely clears or the athlete is referred for further care."³² When the concussion results in loss of consciousness or amnesia lasting longer than 15 minutes, the athlete will require medical treatment to prevent additional complications.³³ Similarly, the onset of drowsiness, paralysis, or language impairment after a concussion is cause for concern about the possibility of more serious complications, usually warranting swift medical examination and medical imaging studies.³⁴ Nonetheless, despite much of what we *do* know about TBIs,

30 See *There's No Such Thing as a Tough Brain*, KING DEVICK TEST, <http://kingdevicktest.com/for-concussions/> (last visited Dec. 2, 2013).

31 See Shepherd, *supra* note 27.

32 Kevin M. Guskiewicz et al., *National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion*, 39 J. ATHLETIC TRAINING 280, 282 (2004).

33 *Id.*

34 See Ropper & Gorson, *supra* note 20, at 168.

many rules about medical treatment (for example, requiring a CT scan for an injured athlete) are based on antiquated science and studies of an age-dissimilar population, which calls into question their generalizability to youth populations.³⁵ Therefore, additional study is required to determine whether current treatment standards are appropriate for youth with TBI symptoms.³⁶

Furthermore, because an initial TBI puts an athlete at greater immediate risk of suffering another, an important part of treatment is prevention of another immediate TBI—specifically, by ensuring that the athlete does not return to competition too quickly following the initial head injury. Although there is rarely unanimity in medical opinion with respect to TBIs, all recommendations for TBI treatment identify removal from potential TBI-inducing activity in the period following an initial TBI as the most critical action to prevent compounding the initial brain injury. During the period immediately following a concussion, a subsequent TBI can have disproportionately severe health consequences, no matter how mild the initial concussion.³⁷ Therefore, as standards have evolved, removing athletes from competition for at least an entire day has become a de facto minimum standard to prevent further aggravation of the initial injury.

Finally, treatment of a TBI must fully account for the array of residual effects of TBIs, which generally include continuation of concussion-related symptoms long after the injury—known as post-concussion syndrome. Post-concussion syndrome is “a constellation of sometimes disabling symptoms, mainly headache, dizziness, and trouble concentrating, in the days and weeks following concussion.”³⁸ These symptoms can last for as little as a few days to as long as one year or more.³⁹ As with other epidemiological matters pertaining to concussions, there is a lack of data from controlled trials for guidance on treatment of post-concussion syndrome, and there is some indication that the extent and duration of some symptoms may be psychosomatic.⁴⁰ Nonetheless, more research is required to more fully understand the scope of the problem.

35 See Micelle J. Haydel et al., *Indications for Computed Tomography in Patients with Minor Head Injury*, 343 NEW ENG. J. MED. 100, 101–02 (2000) (noting that the criteria requiring a CT scan include Glasgow Coma Scale of 15, headache, vomiting, age above 60 years, drug or alcohol intoxication, persistent anterograde amnesia, evidence of traumatic soft-tissue or bone injury above the clavicle, or seizure).

36 Youth were excluded from the original studies, and we now know that youth brains respond differently to concussive events and are still forming—unlike the populations then studied. See Ropper & Gorson, *supra* note 20, at 168.

37 In rare cases, a second impact could cause permanent injury or death.

38 See Ropper & Gorson, *supra* note 20, at 169.

39 *Id.*

40 *Id.*

E. Identification and Treatment of Multiple TBIs

Many professional athletes who suffer MTBIs eventually experience severe health consequences, such as ALS and Chronic Traumatic Encephalopathy (CTE).⁴¹ American football remains the primary focus of CTE research because of post-mortem access to a player's brain tissue, which must be analyzed at the cellular level using specialized proteomics technology to definitively study MTBI consequences.⁴²

MTBIs are considered one cause of CTE, which is a progressive degenerative disease of the brain that has been found in athletes (and others) with a history of repetitive brain trauma.⁴³ Yet the uncertainty about the long-term impact of MTBIs also exists with respect to CTE; scientists are not exactly sure how CTE ultimately manifests.⁴⁴ In addition to this uncertainty, there have not yet been any longitudinal cohort studies that evaluate long-term health outcomes by following athletes with and without TBIs over a multi-decade span.⁴⁵ The absence of such studies makes it impossible to precisely determine the causal relationship, if any, between youth-sports injuries, professional sports injuries, and the subsequent early-onset dementia that has been observed in some former professional athletes.⁴⁶

For many reasons, we simply do not know much about the specific long-term effects of repeated concussions suffered in youth sports. Perhaps one factor in

41 The evolution of science with respect to CTE is fairly remarkable. The discovery began in 2002 by identifying a complex new form of post-concussive dementia, recognized in an autopsy of former NFL player Mike Webster and defined as CTE. Since then, scientific agreement about the phenomenon has coalesced, as CTE was later identified in former Philadelphia Eagles player Andre Waters (who was posthumously diagnosed after committing suicide following numerous concussions), former Houston Oilers linebacker John Grimsley (who died from a self-inflicted gunshot wound in 2008), and many others. Some players, such as Sean Morey, have been diagnosed as having a high likelihood of developing CTE and have retired in an effort to prevent additional harm.

42 See *Field Hearing: Legal Issues Relating to Football Head Injuries Before the H. Comm. on the Judiciary, Part II*, 111th Cong. (2010) (statement of Bennet I. Omalu, Clinical Professor of Psychology, University of California, Davis) [hereinafter *Field Hearing*]. It appears that the Center's public mission and the media visibility of its findings have caused other professional football players to consider brain donation.

43 See, e.g., *What is CTE?*, CTR. FOR THE STUDY OF TRAUMATIC ENCEPHALOPATHY, BOS. UNIV., <http://www-test.bu.edu/cste/about/what-is-cte/> (last visited Dec. 5, 2013).

44 One explanation of CTE development appears as follows: mild trauma to brain causes axons to be sheared, diffuses leakages across membranes, tau formation (possibly in genetically predisposed people). See *Field Hearing*, *supra* note 42.

45 BU's CTE Center has proposed the first such comprehensive study, but its recruitment is still ongoing. See, e.g., *Clinical Studies*, CTR. FOR THE STUDY OF TRAUMATIC ENCEPHALOPATHY, BOS. UNIV., <http://www-test.bu.edu/cste/our-research/clinical-studies/> (last visited Dec. 5, 2013).

46 Kevin M. Guskiwicz et al., *Recurrent Concussion and Risk of Depression in Retired Professional Football Players*, 39 MED. & SCI. IN SPORTS & EXERCISE 903 (2007).

our lack of knowledge is the failure of key constituencies to agree on a system for youth athlete TBI reporting and tracking over time.⁴⁷ A second key reason is that long-term effects can only be truly studied by analyzing brain tissue.⁴⁸ Consequently, athletes cannot be examined for long-term effects during their lifetimes, and those who are examined after death tend to be athletes who played professional sports, particularly football. The athletic careers of professional sports players are obviously very different from those of most youth athletes. It is difficult to make inferences about the impact of MTBIs on young amateur athletes from evidence drawn from professionals.⁴⁹ For these reasons, while there is agreement that MTBIs (especially within a short time) can result in short- and long-term health consequences for youth, there is still much to learn about the scope of the phenomenon and its consequences.⁵⁰

F. Ambiguous Evidence Leads to Ambiguous Policy Logic

As the preceding sections have described, scientific evidence regarding virtually every aspect of the scope and consequences of TBIs is highly underdeveloped. This is especially true with respect to youth involved in school sports. There are no large-scale studies that have measured the impact of public health law interventions in this area. The new state-level legislation discussed in this Article constitutes a type of policymaking driven by news events and studies of unique, individual experiences. Such legislation by anecdote works when empirics are lacking, because their absence creates a vacuum in which any policy seems better than none.⁵¹ While legislating by anecdote has clearly obtained

47 This failure persists for reasons that confound most experts. In October 2013, the Institute of Medicine of the National Academies released the results of its national expert report on youth sports TBIs, concluding that failure to track incidence data was harmful to the goal of reducing TBIs and calling for the Centers for Disease Control and Prevention (CDC) to establish a national surveillance system. See, e.g., *Sports-Related Concussions in Youth: Improving the Science, Changing the Culture*, INST. OF MED. (Oct. 2013), <http://www.iom.edu/~media/Files/Report%20Files/2013/Concussions/concussions-RB.pdf> [hereinafter *Sports-Related Concussions*].

48 There are certainly a variety of physiological methods and metrics by which health professionals can gather the information, but it appears that this more invasive technique objectively reveals more than the other methods.

49 As youth sports TBIs gain wider recognition as a public health issue, observational studies are being conducted by experts to begin collecting data regarding the specific effects of such repeated traumas on children. See, e.g., *Second Concussion Means Longer Recovery Time*, REDORBIT.COM (June 24, 2013), http://www.redorbit.com/news/video/health_2/1112874341/recovery-from-second-concussion-061413/ (showing that youth athletes take longer to recover after their second or third concussion).

50 See Tareg Bey & Brian Ostick, *Second Impact Syndrome*, W.J. EMERGENCY MED. 6, 6–10 (2009).

51 See Theresa Glennon, *Choosing One: Resolving the Epidemic of Multiples in Assisted Reproduction*, 55 VILL. L. REV. 147 (2010).

rapid legislative results in the area of youth sports TBI laws, the value of laws that lack empirical underpinnings remains to be seen.

Despite the potentially problematic absence of empirical evidence, there are two broad areas of agreement about the consequences of youth concussions:

- (1) At a minimum, doctors and others believe that youth are more at risk than adults for short- and long-term TBI effects because the brain is not fully formed until the late teen years moving into adulthood.⁵² As such, experts conclude that injuries to a developing brain are likely to have more long-term consequences (and result in longer healing times) than injuries to a fully-formed brain; and
- (2) Multiple concussions in a short time span likely lead to a more rapid onset of short- and long-term brain dysfunction.

Given these two broad areas of agreement, short-term cohort studies of athletes in the field (of which there are few) attempt to add clarity to the muddled causal links.⁵³ These studies, which describe broad data that can lead to prescriptive measures, seek to provide generic universal recommendations. These generalized recommendations do not differentiate TBI reduction policy prescriptions and minimize the observed differences in TBI rates across diverse populations.⁵⁴ By keeping a broad focus on injury prevention, such studies should point toward a broad public health intervention that encompasses preventing injury, recognizing injury, and caring for injury. However, most of these studies and the legal interventions derived from them focus on prevention of long-term damage from concussions rather than prevention of concussions themselves. Despite this failing, such studies at a minimum provide some broad empirical support for certain types of public health interventions that may positively affect public health with respect to youth populations. The next section discusses two of these broad studies before turning to state-level legislative efforts that have implemented or capitalized on some of the studies' general findings.

⁵² See Bey & Ostick, *supra* note 50, at 6–10.

⁵³ The studies are described in Part II below.

⁵⁴ Although such studies show that incidences and rates can be separated and distinguished by key independent variables such as type of sport, gender, race, and income, policy prescriptions do not attempt to sort TBI reduction strategies by using such inter-group/sport differences to achieve more narrowly tailored TBI reduction goals. See, e.g., P. David Adelson & Patrick M. Kochanek, *Topical Review: Head Injury in Children*, 13 J. CHILD NEUROLOGY 2 (1998) (providing a broad review of some demographic variance indicators).

II. SCIENTIFIC EVIDENCE AND YOUTH SPORTS TBI LAWMAKING

Although no state's youth sports TBI law contains sex- or sport-based policy distinctions, youth concussions do vary substantially on both of these metrics. Few studies have attempted to sort out the causal factors behind these key differences and lay the groundwork for (a) explaining or reducing demographic variance in youth TBIs and (b) influencing how public health laws might reduce these injuries or mitigate their consequences. Instead, policy prescriptions rely on only the broadest of demographic details.⁵⁵

Some recent studies have contributed much to our understanding of how demographic variables might change the incidence, rate, or outcomes of concussions. One such study (the "High School Trends Study") tracked twelve boys' and girls' sports over a multi-year period to analyze risks and trends in youth concussions.⁵⁶ When comparing sports in which both boys and girls participated, the data suggested that girls had a higher rate of concussions than boys did.⁵⁷

However, there is no agreement that sex-based differences in TBI rates exist, and even if there were, it is not clear how such differences should inform youth sports TBI laws. Yet we know that the studies' inclusion of certain gender-specific sports like cheerleading (which is dominated by female athletes) significantly influences data trends and health solutions. When these sports are not included in such studies, it is likely that metrics will focus on comparisons between girls' and boys' divisions of common sports. As such, researchers who highlight the interaction between gender and TBIs would have the lawmaking process focus more on sex-specific youth sports TBI initiatives. These researchers argue against the sex-neutral interventions present in all youth sports TBI laws to date.

Despite evidence that sex-differentiation in TBI rates may exist, TBI research involving youth athletes focuses on areas with few to no cross-sex comparisons. These studies use different methods, lead to different conclusions, and sometimes contradict each other. For example, one recent high school sports study concluded that, holding other factors constant, there were no sex differences in the severity of concussions or short-term outcome measures, like recovery time.⁵⁸ Some studies have found that female athletes have higher overall TBI rates,

⁵⁵ See Gessel, *supra* note 23.

⁵⁶ See Andrew E. Lincoln et al., *Trends in Concussion Incidence in High School Sports: A Prospective 11-Year Study*, 39 AM. J. SPORTS MED. 958 (2011) [hereinafter High School Trends Study].

⁵⁷ *Id.*

⁵⁸ See Leah J. Frommer et al., *Sex Differences in Concussion Symptoms of High School Athletes*, 46 J. ATHLETIC TRAINING 76 (2011).

while others suggest that such data may be due to heightened reporting of female TBIs. Broadly speaking, one cause of the heightened reporting could be a cultural bias in favor of being more protective of female athletes than their male counterparts.⁵⁹ The same cultural biases may encourage male athletes to play even when they exhibit signs of concussions. Some boys suffering from head injuries may not report their symptoms for fear of being removed from play, which would depress male reporting even though overall rates might remain high.⁶⁰

While such available research might lean toward a more sex-nuanced approach to youth sports TBI laws,⁶¹ other evidence points toward a sex-neutral public health approach to TBIs. But there is one area of general agreement across the sex-based studies: boys' football dominates cases of TBIs, accounting for more than half of all reported youth TBIs.⁶² This rhetorical and scientific focus by commentators and policymakers on one particular sport's tendency to produce TBI events, in conjunction with existing anecdotal evidence of negative long-term consequences on professional football athletes, may be the motivating factor behind the football-based narratives underlying most (if not all) legislative discussions of youth sports TBI laws. On the other hand, such a focus in policymaking discussions could also indicate the NFL's influence on youth sports concussion policymaking.

Other studies suffer from similar data deficiencies when trying to parse population or incidence differences. For example, most studies of TBIs in youth sports do not account for non-scholastic athletes (athletes who play in recreational leagues or informal non-school community events), who tend to be younger and are often at even greater risk of concussion than youth athletes who play in school-sanctioned sports. Likewise, the data presented in such studies, by and large, does not differentiate between concussed athletes based on their socioeconomic or demographic information.⁶³ Finally, while it is clear that those playing high-school football are the most susceptible to concussions, more study could provide information on whether athletes with certain physical

59 See, e.g., PATRICIA A. VERTINSKY, *THE ETERNALLY WOUNDED WOMAN: WOMEN, DOCTORS, AND EXERCISE IN THE LATE NINETEENTH CENTURY* 1–7 (1994).

60 See Mark R. Lovell et al., *Inaccuracy of Symptom Reporting Following Concussion in Athletes*, 34 MED. & SCI. SPORTS & EXERCISE, at S-298 (2002).

61 For example, calibrating return-to-play guidelines by sex if evidence shows that girls' injuries take longer to heal than boys' injuries, or calibrating symptom differentiation by sex to take into account whether boys and girls have higher baseline rates of common symptoms (like headaches) prior to concussions.

62 See Lincoln et al., *supra* note 56.

63 See Alan Schwarz, *As Injuries Rise, Scant Oversight of Helmet Safety*, N.Y. TIMES, Oct. 20, 2010, <http://www.nytimes.com/2010/10/21/sports/football/21helmets.html> (describing fiscal constraints as one factor that school districts face when updating clearly outdated equipment intended to reduce injury risks).

characteristics are more vulnerable to concussions.

Most studies that have analyzed youth populations have produced inconsistent results. However, several attempts have been made to apply a more rigorous analysis of the causes and consequences of sports-related brain injuries at the collegiate level, which can serve as a useful case study for youth athletics. The NCAA's comprehensive analysis ranks as one of the most authoritative studies of college sports to date (the "NCAA Study").⁶⁴ The NCAA Study focused on the recovery period after an athlete's initial injury and monitored both recovery time and the effects on an individual player following an on-field concussion. The study sought to bring a stronger empirical focus to determining when it is safe for a player to return to the game and the times during which a post-concussive player is most vulnerable to re-injury.

The NCAA Study evaluated the effects of sports-related concussions from the 1999-2001 football seasons. The NCAA studied 1631 football players from 15 NCAA Division I, II, and III institutions.⁶⁵ Ninety-four players who were identified as having a football-related concussion were enrolled in an extensive injury-assessment protocol, and 70 of them completed the 90-day protocol. The study's controls included a pre-season baseline evaluation with a health history questionnaire and a non-injured control team member. Players who suffered from a concussion⁶⁶ were tested immediately after the injury, two to three hours after, and again on days 1, 2, 3, 5, 7, and 90 post injury. The athletes underwent neuropsychology testing prior to the injury, as well as post injury on days 2, 7, and 90.

The study revealed that concussions accounted for a significant percentage of total athlete injuries in a number of college sports: ice hockey (12.2%), football (8%), and soccer (4.8%) in particular. Given the large number of program participants, college football had the highest overall number of brain injuries per year. The study stated that between 3 and 8% of hockey and football players sustain a concussion in each season, and the trends showed an increase during the years before the study's conclusion.

The study's most important finding was that by seven days after the injury,

64 See Michael McCrea et al., *Acute Effects and Recovery Time Following Concussion in Collegiate Football Players: The NCAA Concussion Study*, 290 JAMA 2556 (2003). The paragraphs that follow all reference this study.

65 The study was conducted before the recalibration and renaming of the divisions.

66 The study used the American Academy of Neurology Guideline for Management of Sports Concussion definition: "An injury resulting from a blow to the head causing an alteration in mental status and 1 or more of the following signs: Headache, nausea, vomiting, dizziness/balance problems, fatigue, difficulty sleeping, drowsiness, sensitivity to light or noise, blurred vision, memory difficulty, and difficulty concentrating." American Academy of Neurology, *Practice Parameter: The Management of Concussion in Sports (Summary Statement): Report of the Quality Standards Committee*, 48 NEUROLOGY 581 (1997).

91% of players had full resolution of symptoms. During that seven-day period, cognitive and verbal impairment persisted through at least two days, and milder cognitive effects persisted through five but often as long as seven days. Athletes had acute difficulties achieving balance during the 24 hours after injury but appeared to resolve these issues by the fifth day following the injury.

The NCAA Study also reinforced some now-conventional wisdom. The Study concluded that concussions can affect multiple domains (ionic, metabolic, and physiological) and can adversely affect core cerebral function for several days to weeks. In addition, the trajectories of recovery for these domains are not aligned. Symptoms, cognition, and balance recover at different speeds, complicating the process of determining when a player is safe to return to the field.

The NCAA Study's authors offered a few initial conclusions and recommendations. First, even after 90 days, players who experienced a concussion performed less well than control players did on verbal fluency measures. Second, college football players required, at a minimum, several days to recover after sports-related concussions. Third, cerebral dysfunction occurred in cases without typical indicators (like loss of consciousness and posttraumatic amnesia). Therefore, given the consequences of missing these subtle indications, medical professionals and standards-creating bodies should consider expanding or changing traditional definitions of concussion. Despite these strengths, the study failed to capture players who had the most "mild" forms of concussion, and the study's design was not fully able to identify factors that predict recovery across all areas or an appropriate duration of symptom-free recovery that would minimize risks of re-entry onto the playing field.

The High School Trends Study and the NCAA Study attempted to combine prior epidemiological studies into testing instruments and then evaluate extant knowledge against a sample of existing populations. Both studies, however, engaged in information gathering as much as hypothesis testing. To the extent that these two studies advanced prescriptive understandings of how regulatory bodies should approach incidences of TBIs, conclusions were sparse but clear: (a) TBIs are most prevalent in football with respect to overall incidences of TBIs, and a football-specific approach likely would achieve the most efficacious results; (b) incidences of low-grade concussions were much higher than previously estimated; (c) instances of low-grade concussions were fairly difficult to detect using traditional methods but easier to identify using baseline methodologies; and (d) instances of full recovery from TBIs took substantially longer than 24 hours.

Based upon these two prominent recent studies of TBIs in youth sports populations, one might expect that public health TBI interventions would: (a) focus on reducing the direct instances of TBIs, particularly with sport-specific solutions; (b) provide information-gathering and reporting mechanisms on the

overall number of TBIs; (c) provide rigorous methods by which athletes, parents, and schools can detect subtle changes in athlete physiological characteristics; and (d) substantially extend recovery times by requiring lengthy “time-out” windows of five days or more following an athlete’s suspected incurrence of a TBI. But the public health law interventions did not do any of the above. To understand how the policy prescription diverged from the medical evidence, the following section analyzes the component parts and provisions of all fifty youth sports TBI laws.

III. RUSHING THE FIELD: YOUTH SPORTS TBI LAWMAKING (2009-2012)

A study of the entire set of state youth sports TBI laws is a complex endeavor given the many causal inputs and policy variations. To provide the broadest and most accurate framing, this Part proceeds as follows. First, in Section III.A, the galvanizing incident prompting such laws will be described in its historical context. In Section III.B, a description of data gathering and methods will explain how these laws were analyzed and separated into their component parts. Next, Section III.C will examine common themes across states. Section III.D will analyze provisions that are only present in a minority of states but can be classified as “Lite-Experimental Policy” additions to a comprehensive youth sports TBI law. Section III.E will examine unique provisions in state laws that are of interest for further refinement of youth concussion laws and policies. Concluding the analysis, Section III.F will explore the common causes of policymaking choices during the legislative process, focusing on the role of the National Football League. The cumulative theory guiding each of these sections is that the NFL, through its unique role as a dominant interest group, established the content of states’ youth sports TBI laws. The NFL’s vigorous advocacy caused state legislatures to act swiftly, which minimized the role of scientific evidence and policy experimentation in the youth sports TBI lawmaking process.

A. *The “Galvanizing Incident” Model of Policymaking*

In 2006, Zackery Lystedt, then a 13-year-old middle-school student, sustained a severe head injury during a football game⁶⁷ after he returned to play shortly after sustaining a TBI that was not properly evaluated.⁶⁸ This post-TBI play led to additional head injury and eventually caused permanent brain

67 See *Encouraging Signs for Teen Seriously Injured During Football Game*, KOMONEWS.COM, Oct. 17, 2006, <http://www.komonews.com/news/local/4422577.html>.

68 Matt Rybaltowski, *Young Player Helps Turn Trauma into Action on Concussions*, CBSSPORTS.COM, Feb. 14, 2010, <http://www.cbssports.com/nfl/story/12928497/young-player-helps-turn-trauma-into-action-on-concussions>.

damage.⁶⁹ Over the next several years, the Brain Injury Association of Washington worked with Lystedt's family to build a coalition of community, business, and sports organizations.⁷⁰ This effort ultimately led the Washington State Legislature to unanimously pass the 2009 Zackery Lystedt Law,⁷¹ which set guidelines and standards to help recognize concussions and protect young athletes from immediate re-injury after sustaining an initial TBI.⁷² The statute was the first of its kind, specifically addressing TBIs in youth sports at the state level.⁷³

To reduce concussion-related injuries among youth athletes, the Washington State law combined three essential elements that would become the basis for similar laws in other states.⁷⁴ (1) The law required athletes, parents, and coaches to receive annual education sessions about recognizing the symptoms of concussions. If a young athlete is suspected of having a concussion, (2) he or she would be removed from a game or practice, and (3) would not be permitted to return to play without clearance from a licensed healthcare professional.⁷⁵ These provisions promoted injury recognition and post-injury care, two laudable public health goals. Nonetheless, they fell short of the full protections suggested by the epidemiological and empirical studies of TBIs discussed in Parts I and II of this Article. Most notably, the Zackery Lystedt Law did not try to prevent initial concussions within a public health law intervention framework. Further complicating matters, the Lystedt law defined the problem in a fairly tight bandwidth—the problem was “re-injury” and so the solutions focused attention on secondary prevention measures and avoided engaging primary prevention strategies as a legislative tool.

Following the enactment of the Zackery Lystedt Law, the NFL supported and engaged the issue by defining the problem narrowly. The NFL lobbied to have other states as well as the United States Congress follow suit and enact similar—indeed, almost identical laws to the Lystedt Law.⁷⁶ To advance this goal, NFL

⁶⁹ *Id.*

⁷⁰ *Id.*

⁷¹ *Id.*

⁷² *The Zackery Lystedt Law*, NFLEVOLUTION.COM, Aug. 9, 2012, <http://www.nflevolution.com/article/The-Zackery-Lystedt-Law?ref=270> [hereinafter *Zackery Lystedt Law*].

⁷³ This rapid response to one galvanizing incident may be part of a larger phenomenon of legislating by anecdote. See Glennon, *supra* note 51, at 149 n.13 (“Legislative responses to highly unusual but extremely salient events often address issues of immediate public concern but typically ignore larger or more common structural issues.”). Other examples include the Megan’s Laws passed by various states, see, e.g., N.J. STAT. ANN. 2C:7-2, as well as the Terry Schiavo Law, Act for the Relief of the Parents of Theresa Marie Schiavo, Pub. L. No. 109-3, 119 Stat. 15 (2005).

⁷⁴ Zackery Lystedt Law, *supra* note 72.

⁷⁵ *Id.*

⁷⁶ *Id.*; see H.R. Res. 6172, 111th Cong. (2010).

Commissioner Roger Goodell sent a letter to 44 states' governors urging legislation.⁷⁷ Goodell also encouraged political and athletic leaders to help raise concussion awareness and promote proper post-concussive treatment.⁷⁸ The NFL urged states to adopt the same three core elements as the Washington law, calling them the "three tenets of model legislation."⁷⁹ However, during the process of encouraging the adoption of model legislation, the first tenet requiring parents and coaches to be annually educated "about the dangers of concussions"⁸⁰ evolved to provide instead that parents or guardians of young athletes simply "sign a concussion information form" prior to allowing their child to engage in youth sports.⁸¹

The focus of legislative efforts on a more narrowly defined problem, following passage of the Lystedt Law was shaped, in part, by the NFL's early and visible involvement. Given this proactive effort by an interested and influential private for-profit interest group, it is not surprising that subsequent TBI legislation in many states exhibited remarkable uniformity based on the NFL's suggestions.

Defining the problem as re-injury allowed, as such definitions often do, for a consistent message and a default shaping of policy prescriptions.⁸² The role that the NFL played in promoting the Zackery Lystedt Law to other states as the basis for strikingly similar, if not identical, legislation can be viewed as a form of regulatory policy capture that effectively shaped the policy discussion and dwarfed policy input from other sources.⁸³ While the laws will likely have some positive impact, the lawmaking process that created them may have unintentionally crowded out more innovative strategies. Although the scientific and analytic background for youth TBI policymaking is currently lacking, youth athletes are undoubtedly a population that deserves thoughtful, empirically grounded safety regulations that are rooted in a more expansive definition of the problem than allowed by the messaging of the NFL and others.

77 *Concussion Legislation by State*, NFLEvolution.COM, <http://www.nflevolution.com/article/Concussion-Legislation-by-State?ref=767> (last visited Feb. 9, 2013) [hereinafter *Concussion Legislation*]; see, e.g., Letter from Roger Goodell, Commissioner, Nat'l Football League, to Christopher Christie, Governor, N.J. (May 21, 2010), <http://nflevolution.com/wordpress/wp-content/uploads/2012/08/NJ-Gov-letter.pdf>.

78 *Id.*

79 *Concussion Legislation*, *supra* note 77.

80 See Letter from Roger Goodell, *supra* note 77.

81 *Concussion Legislation*, *supra* note 77.

82 See generally JOHN W. KINGDON, AGENDAS, ALTERNATIVES, AND PUBLIC POLICIES (2003).

83 See *infra* notes 123–124 and accompanying text.

B. *Evaluating Youth Sports TBI Laws—Data and Methodology*

The current state of the field examined in this Article includes the youth sports TBI laws enacted by 49 states and the District of Columbia⁸⁴ as of October 1, 2013. The state-by-state examination includes only the text of the law and does not include other provisions that might require additional actions of coaches or school districts, such as state agency administrative rules or district level policies. To facilitate future empirical evaluation of these laws, the legal research was conducted in accordance with scientific principles of transparency and reproducibility⁸⁵ and is embodied in an open source protocol, codebook, and dataset available online at the Public Health Law Research Policy Surveillance Web Portal.⁸⁶

The results of the legislative analysis are clear, identifying a high level of uniformity across state policies. This policy uniformity can best be assessed by comparing the clusters of common policy agreement among the laws and with reference to the Lystedt Law discussed earlier.

C. *A Uniform Policy Solution?*

When faced with a large public health problem, there are numerous policy options that might be employed with varying degrees of effectiveness. But with respect to TBIs in youth populations, the questions begin at a more basic level. What precisely is the problem?⁸⁷ And why is public health law the place to solve it? Although there is consensus about the scope and cause of TBIs in adult populations, researchers, scientists, and advocates do not agree about the magnitude of youth TBIs or their root causes.⁸⁸

Given a problem of such undetermined scope, whether public health law (rules, regulations, or other mandates) could influence an optimal public health

84 Mississippi is the only state without a youth sports TBI law as of October 1, 2013.

85 See, e.g., Jennifer K. Ibrahim et al., *State Laws Restricting Driver Use of Mobile Communications Devices: Distracted Driving Provisions, 1992-2010*, 40 AM. J. PREVENTIVE MED. 659, 660–61 (2011) (describing a similar comprehensive statutory study).

86 See LAWATLAS, <http://lawatlas.org> (last visited Dec. 3, 2013).

87 If the problem were defined as youth having too many TBIs, preventive measures would likely be more invasive (e.g., banning certain sports or activities). If the problem is schools failing to enact simple risk-reduction measures, law might provide for simple mandates. Or, if the problem is that cultural norms prevent more invasive sport-specific measures, perhaps the solution might involve education about sport-specific dangers in an attempt to proactively shape public opinion.

88 See, e.g., Nat'l Ctr. for Injury Prevention & Control, *Traumatic Brain Injury in the United States: Assessing Outcomes in Children*, CENTERS FOR DISEASE CONTROL & PREVENTION (2000), http://www.cdc.gov/traumaticbraininjury/pdf/TBI_assessing.pdf (noting that “TBI is often described as the leading cause of disability in children, but data to support this assertion are lacking Currently no population-based studies of the outcomes of TBI among children and youth exist to provide national estimates of TBI-related disability and document the need for services”).

outcome (reducing youth sports TBIs and their health consequences) was not clear at the outset of the proposed policy solutions.⁸⁹ Therefore, although public health advocates and health professionals have long known that TBIs and MTBIs may lead to short- and long-term health consequences,⁹⁰ no public health law interventions were developed in response to the growing awareness of such problems in youth populations prior to 2009. However, the confluence of the events described earlier, a growing consciousness of health risks among college and professional athletes, numerous tragic professional athlete deaths, and the motivation of a powerful outside actor (the NFL) created an environment ripe for state legislative action.⁹¹

Between 2009 and 2013, 49 states and the District of Columbia enacted one or more youth sports TBI laws. There are no states that have banned traditional youth sports with high TBI risks or that have set out legal regimes attempting to govern particular sports techniques by legislation or regulatory oversight.

In concert with the public advocacy of the NFL⁹² and with the inclusion of the three tenets of Washington's Lystedt Law,⁹³ many of the other states' legislation includes all three tenets in one form or another. As previously stated, the laws require: (1) that a youth athlete who appears to have suffered a concussion be removed from play or practice at the time of the suspected concussion for a minimum of 24 hours;⁹⁴ (2) that a youth athlete be cleared by a

89 Some intervention, of course, would coalesce on the minimum that we do know, such as the consensus that it is generally harmful to have multiple TBIs and certainly within a short period. Nonetheless, if the overall optimal goal is to generally reduce TBIs in the first instance, the epidemiology is less clear.

90 Various advocacy and trade groups set out detailed concussion prevention proposals many years ago, and the CDC has provided detailed information and guidelines for more than a decade. See, e.g., *Traumatic Brain Injury: A Case for Prevention*, MASS. TRAUMATIC BRAIN INJURY PREVENTION TASK FORCE (2007), <http://www.mass.gov/cohhs/docs/dph/com-health/injury/tbi-case-prevention.pdf> (report convened by the Massachusetts Department of Public Health detailing existing state TBI data and recommending prevention strategies, including surveillance and evaluation, policy and enforcement, and education and training).

91 See Alan Schwarz, *N.F.L. Players Shaken by Duerson's Suicide Message*, N.Y. TIMES, Feb. 20, 2011, <http://www.nytimes.com/2011/02/21/sports/football/21duerson.html>.

92 See *Concussion Legislation*, *supra* note 77.

93 See Zackery Lystedt Law, WASH. REV. CODE § 28A.600.190 (2013).

94 But there is not scientific agreement that the 24-hour return-to-play minimum is optimal (as opposed to a longer minimum rest), nor is there agreement for how medical professionals might design an optimal period of recovery. See Paul McCrory et al., *Consensus Statement on Concussion in Sport: The 4th International Conference on Concussion in Sport Held in Zurich, November 2012*, 47 BRIT. J. SPORTS MED. 250 (2013). It may be that the focus on a 24-hour minimum in the NFL's messaging, contrary to most scientific studies, is one example of the impact such messaging had in the universal adoption of the 24-hour minimum as opposed to a longer, more supported waiting period.

licensed health care professional trained in the evaluation and management of concussions before returning to play or practice;⁹⁵ and (3) that each year, athletes, parents, and coaches receive education or a information sheet about recognizing the symptoms of concussions.⁹⁶ Thus, key features across all youth sports TBI laws include a focus on *secondary*, not *primary* prevention, as well as a general adherence to the Lystedt framework with minimal policy experimentation.⁹⁷ In Table 1 on the following page, state laws are compared by reference to their most common elements and those of the Lystedt framework.⁹⁸

95 See *Concussion Legislation*, *supra* note 77.

96 See *id.*; Letter from Roger Goodell, *supra* note 77.

97 There are other explanations for states following the Lystedt model (and subsequent state frameworks based on this model). A behavioral critique of lawmakers might suggest that such lawmakers often follow the status quo or default policy position on public health issues and take the path of least resistance when evaluating a new policy measure in the face of an existing public health policy already promulgated in another jurisdiction. See, e.g., Eric J. Johnson & Daniel Goldstein, *Do Defaults Save Lives?*, 302 SCIENCE 1338 (2003).

98 This table does not represent all of the variables, and statutes that have been modified since enactment are not fully incorporated.

REFEREEING THE PUBLIC HEALTH

Table 1. Key Features of Youth Sports TBI Laws (2009 - 2013)

STATE	Has Youth Sports TBI Law	Requires Removal From Play	Requires Parental Notification	Return to Play Requirements					Training	TBI Information/Awareness Requirements				
				Specifies Return To Play	Requires Clearance by a Health Professional	Requires Written Clearance	Health Professional Trained in TBI Management Required	Clearance by a Medical Doctor Required		Requires Special Training for Coaches	Requires Information Sheet Distribution	Parent's Signature Required	Student's Signature Required	Information Sheet Distributed on an Annual Basis
AK	✓	✓		✓	✓	✓	✓				✓		✓	✓
AL	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓	✓
AR	✓	✓		✓	✓	✓	✓				✓		✓	✓
AZ	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
CA	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
CO	✓	✓	✓	✓	✓	✓	✓		✓					
CT	✓	✓		✓	✓	✓	✓		✓					
DC	✓	✓		✓	✓	✓					✓	✓	✓	
DE	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓	✓
FL	✓	✓		✓	✓	✓	✓					✓		✓
GA	✓	✓		✓	✓	✓	✓				✓			✓
HI	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
IA	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
ID	✓	✓		✓	✓		✓				✓			✓
IL*	✓													
IN	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
KS	✓	✓		✓	✓	✓		✓			✓	✓	✓	✓
KY	✓	✓		✓	✓	✓	✓			✓				
LA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
MA	✓	✓		✓	✓	✓	✓			✓	✓			
MD	✓	✓		✓	✓	✓	✓				✓	✓	✓	
ME*	✓													
MI	✓	✓		✓	✓	✓	✓				✓	✓	✓	
MN	✓	✓		✓	✓	✓	✓		✓					
MO	✓	✓		✓	✓	✓	✓				✓	✓		✓
MS**														
MT	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
NC	✓	✓		✓	✓	✓					✓	✓	✓	✓
ND	✓	✓		✓	✓	✓	✓			✓	✓			
NE	✓	✓	✓	✓	✓	✓	✓				✓			✓
NH	✓	✓	✓	✓	✓	✓	✓							
NJ	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NM	✓	✓		✓	✓					✓	✓	✓	✓	✓
NV	✓	✓		✓	✓	✓					✓	✓	✓	✓
NY	✓	✓		✓	✓	✓		✓	✓	✓	✓			
OH	✓	✓		✓	✓	✓			✓	✓	✓	✓		✓
OK	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
OR	✓	✓		✓	✓	✓				✓				
PA	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
RI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
SC	✓	✓		✓	✓	✓		✓			✓	✓		✓
SD	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
TN	✓	✓		✓	✓	✓			✓	✓	✓			✓
TX	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
UT	✓	✓		✓	✓	✓	✓				✓	✓		
VA	✓	✓		✓	✓	✓					✓	✓	✓	✓
VT	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
WA	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
WI	✓	✓		✓	✓	✓	✓				✓	✓	✓	✓
WV	✓	✓		✓	✓	✓	✓			✓	✓	✓	✓	✓
WY	✓								✓					
Total	50	47	7	47	47	44	33	8	25	41	32	28	34	

Most states follow the first Lystedt tenet—removal from play following an actual or suspected TBI. Forty-six states and the District of Columbia have such a requirement, and all but Wisconsin and Ohio require removal for a minimum of 24 hours. The states that have such a minimum time length removal requirement do not substantively vary in their approach to re-entry. No state law mandates a minimum removal time beyond 24 hours, although the accompanying health evaluation and clearance provisions (discussed below) may effectively create a longer minimum removal provision by default.

The states differ much more on the second tenet of the model legislation, which requires that athletes with concussions or TBIs be cleared by a medical or health professional before return to play. Forty-four of the 50 jurisdictions with youth sports TBI laws have provisions requiring written clearance, but the type of medical professional required to conduct the evaluation varies widely by state. Only eight states require that a medical doctor provide the clearance. Some require that such professionals must be medical specialists of some kind, such as licensed physicians (with or without TBI training). Still other states permit neuropsychologists to provide clearance. Athletic trainers and nurses (who can also be specialists in many areas) are included in some clearance provisions but not others. For example, some suggest that athletic trainers could have a conflict of interest, especially in instances where the player's return to the active roster would benefit the team. However, whether this potential conflict of interest influences clearance decisions in practice is unknown.⁹⁹ Perhaps responding to ambivalence about laws that allow non-specialist physicians to provide clearance, Rhode Island eventually introduced legislation to amend its youth concussion policy to require clearance from a licensed healthcare professional "trained in the evaluation and management of concussions" rather than by a physician.¹⁰⁰ This focus on all providers having TBI training eventually emerged as a new majority position. Thirty-two jurisdictions require that a clearance provider be specifically trained in TBI recognition and/or symptom management. Why the remaining jurisdictions do not is unclear. But generally speaking, this example of divergence in policymaking with respect to one portion of one provision

99 See, e.g., Press Release, Am. Acad. of Pediatrics, High Schools with Athletic Trainers Have More Diagnosed Concussions, Fewer Overall Injuries (Oct. 22, 2012), <http://www.aap.org/en-us/about-the-aap/aap-press-room/pages/High-Schools-with-Athletic-Trainers-have-More-Diagnosed-Concussions-Fewer-Overall-Injuries.aspx>.

100 R.I. GEN. LAWS § 16-91-3 (2013); Telephone Interview with Rep. Raymond E. Gallison Jr., Rhode Island House of Representatives (June 20, 2011) (noting that Rhode Island Interscholastic League, through its sports medicine subcommittee, brought forth recommendations to Rep. Gallison advising that athletes who suffered or were suspected of suffering concussions should only be evaluated by specialists in concussion management).

demonstrates a lack of general scientific (and political) consensus about why a particular health professional should perform this clearance function or whether one type of professional is preferable to another.

With respect to the third tenet, 41 states and the District of Columbia require that some form of TBI information be distributed to parents and student athletes, yet only 34 of these jurisdictions require distribution on an annual basis. Parents or guardians must provide a signature release that they have received the information. The substance of such education and the form's language are not specified in the laws, although the Centers for Disease Control and Prevention's (CDC) well-disseminated materials have been explicitly mentioned in some laws as an initial guidepost for creating such educational materials.

Many states have innovated beyond the three tenets while staying well within their spirit. Twenty-five jurisdictions require additional coach education for recognizing the symptoms of TBIs in youth sports. However, the additional education requirements substantially vary. Some states, such as Vermont and New York, require coaches to receive training every two years,¹⁰¹ while Arkansas requires that coaches receive training every three years.¹⁰² Although a few states, such as Michigan, provide general guidelines regarding the broad contours of a training program or methods of distributing training materials, none of the laws studied here define exactly how such coach education efforts should be constructed. The laws delegate the task of determining the educational programs' content to third parties such as school boards, athletic associations, or departments of health. However, in many of these states, coaches may be subject to concussion education requirements mandated by their voluntary membership associations or athletic governing bodies, which can be more rigorous than the state-mandated requirements. There is no compelling reason why states should avoid training requirements in youth sports TBI laws, given the wide availability of information and its low cost.¹⁰³

States also share a common understanding that the implementation of their youth sports TBI laws should be handled collaboratively with key stakeholders. Most states designate one or several statewide groups to develop policies and standards for youth concussion awareness, as well as protocols that may be adopted by individual school districts and athletic programs. The groups creating the guidelines are often the state departments of education and/or health;¹⁰⁴ however, some states require other types of organizations, both public and

101 See N.Y. EDUC. LAW § 305 (McKinney 2013); VT. STAT. ANN. tit. 16, § 563 (2013).

102 See ARK. CODE ANN. § 6-18-708 (2013).

103 See, e.g., *Concussion in Sports: What You Need to Know*, NFHS LEARNING CENTER, <http://www.nfhslearn.com/electiveDetail.aspx?courseID=38000> (last visited Dec. 5, 2013).

104 See, e.g., MD. CODE ANN., EDUC. § 7-433 (West 2013); N.Y. EDUC. LAW § 305 (McKinney 2013); 24 PA. CONS. STAT. § 5323 (2013).

private, to coordinate and develop youth concussion guidelines. For example, Idaho requires the State Board of Education to work in collaboration with the Idaho School Activities Association to “provide access to appropriate guidelines and information that identify the signs and symptoms of a concussion and head injury.”¹⁰⁵ Similarly, South Dakota specifies that the state’s high school athletic conference and its Department of Education “shall develop guidelines to inform and educate member schools, coaches, athletes, and the parents or guardians of athletes, of the nature and risk of concussion.”¹⁰⁶ Presumably, these guidelines are created by those who have the resources and expertise to provide local districts with scientifically accurate and current information. This type of localized information gathering and policy specification could be more efficient compared to a requirement that each district create separate, and perhaps incomplete, youth concussion policies.

This policy innovation with respect to educating key constituents (e.g., coaches, parents, youth athletes, schools) comes with some scientific uncertainty about the link between constituent education and the policy goal of such education—reducing youth sports TBIs. For example, although the CDC has created broad youth sports TBI guidance, the fidelity of state and local informational materials to CDC guidance is currently unknown.

Therefore, the particulars of the coach and parent “education” components of these laws present a number of important issues for implementers and researchers. First, the efficacy of education and consent in helping parents or coaches prevent, identify, and respond to TBIs in this context is unknown. Second, the content of the required education is in most cases not specified in the laws—thus allowing for wide variance in implementation. Third, there is divergence between the education regimes required for key stakeholders (e.g., parents vs. coaches), which may produce inconsistent responses to potential TBI events. Fourth, as some training materials already exist, a system in which each state crafts its own standards from scratch may be inefficient. Finally, some coaches may already receive substantial training through membership athletic associations and other non-government entities, but the laws do not contemplate the integration of such education into a more uniform education approach among stakeholders such as coaches, parents, athletes, and schools.

Overall, the above concerns are examples of the limitations of relying on the Lystedt’s limited framework. The Lystedt framework, which does not provide detailed guidance for how coaches and parents should be educated, thus allows for experimentation and ambiguity at the expense of policymaking clarity through a centralized framework. Further, because the Lystedt framework provides only general guidance, states could have pushed beyond its provisions

105 See IDAHO CODE ANN. § 33-1625 (2013).

106 S.D. CODIFIED LAWS § 13-36-9 (West 2013).

when creating their own laws. Some states did take such an incrementally experimental approach in an effort to more aggressively solve the problem and innovate outside of the Lystedt framework, and to such experimentation we now turn.

D. The Role of Policy Experimentation (Purposeful and Accidental)

One example of a policy innovation adopted by a small minority of states (and likely consistent with sound science) is the requirement that states review and update their state youth concussion information sheets to determine the program's effectiveness and make revisions as new techniques to identify and treat concussions become available. Without evaluative metrics and mandatory state-level reporting built into the framework, however, how will the states know how to update their policies? If the states with annual review criteria have no way to analyze results from the field or assess whether the legislation is producing its desired effect, it will be difficult to successfully apply a logic/feedback model to improve existing information.

Defining the scope of "athletic activity" may have resulted in a form of accidental policy experimentation. Many states failed to broadly define youth athletic activity, thus narrowing the scope of the youth sports covered under existing laws. Specifically, the majority of youth sports TBI laws do not include non-scholastic athletic team activity, such as athletic clubs and leagues that are not affiliated with a school or school district. In addition, Utah is one of the handful of states where the youth concussion policy specifically includes physical education,¹⁰⁷ an athletic activity in which more than half of high-school students in the United States participate.¹⁰⁸ Furthermore, few chose to include programs or teams associated with county recreation departments in their youth concussion guidelines. While there may be a higher prevalence of concussions among athletes participating in organized scholastic sports such as football, lacrosse, or soccer,¹⁰⁹ this does not mean that children participating in other activities such as physical education class could not benefit from coverage by provisions in youth sports TBI legislation.¹¹⁰ The states that have not yet adopted

107 UTAH CODE ANN. § 26-53-401 (West 2013).

108 Ctr. for Disease Control & Prevention, *Youth Risk Behavior Surveillance—United States, 2011*, MORBIDITY & MORTALITY WKLY. REP., June 8, 2012, at 37.

109 See Lincoln et al., *supra* note 56. Football accounts for more than half of concussions, while baseball had the lowest incidence of concussion. In girls' sports, soccer represented the highest proportion of concussions, followed by cheerleading, basketball, and lacrosse. Overall concussion rates for boys are more than double those for girls. *Id.*

110 *Youth Concussion FAQ*, NFL, <http://www.nfl.com/news/story/0ap1000000228346/article/youth-concussion-faq> (last updated Aug. 9, 2013) ("Concussions can occur in athletes of any age and in any sport or recreational activity. In fact, each year, U.S. emergency departments treat

such policies (and even those that have) could consider broadening their definition of athletic activity to be more inclusive.

Nonetheless, even if these laws do not expansively cover all forms of youth athletics, during the implementation of youth concussion policies, individual school districts could voluntarily apply concussion training to physical education courses and students. Some states even have language in their youth sports TBI laws encouraging youth sports programs not included in the legislation to adopt a similar policy¹¹¹ or requiring those organizations that use school facilities to comply with the youth concussion policy.¹¹² While it is certainly possible that schools or youth concussion policymakers may be able to apply safety policies to athletic activities not covered by the law, the lack of a specific provision in such laws means that children participating in non-scholastic athletic teams may receive a lower standard of care if they suffer a TBI.

E. Substantive Policy Innovation: Evidence, Theory, and Practice

In addition to the three tenets discussed in Section III.C and other policy experimentations discussed in Section III.D, a few states have developed novel provisions that could be the basis for future model legislation. The role of baseline testing provides one such area for experimentation. Rhode Island's youth sports TBI law previously contained a provision that "[s]chool districts are encouraged to have all student athletes baseline or [ImPACT] tested prior to the start of every sport season."¹¹³ Subsequently, Rhode Island passed legislation to remove the ImPACT testing requirement and substitute the test with a free, non-proprietary evaluation tool.¹¹⁴ This type of provision could prove to be more effective in actually detecting and treating TBIs among young athletes than simple TBI identification training. However, the administration of such a test for large groups of young athletes may have an added cost for school districts and parents.

an estimated 135,000 sports-related and recreation-related TBIs, [i]ncluding concussions, among children ages 5 to 18.").

111 See, e.g., R.I. GEN. LAWS § 16-91-4 (2013) ("All other youth sports programs not specifically addressed by this statute are encouraged to follow the guidance set forth in this statute for all program participants who are age nineteen (19) and younger.").

112 See ARIZ. REV. STAT. ANN. § 15-341 (2013) ("A group or organization that uses property or facilities owned or operated by a school district for athletic activities shall comply with the requirements of [the youth concussion policy]").

113 R.I. GEN. LAWS § 16-91-3 (2010).

114 See 2010 R.I. Pub. Laws. 112 (codified as amended at R.I. GEN. LAWS ANN. § 16-91-3 (West 2013)); Telephone Interview with Rep. Raymond E. Gallison Jr., Rhode Island House of Representatives (June 20, 2011) (stating that the Rhode Island Interscholastic League, the agency to which all high-school athletic teams belong, sought to remove the baseline testing provision of the law since it was a proprietary test; at the time of the interview, a substitute baseline testing was being developed by subcommittee for free use).

The NCAA Study and many others have concluded that proper baseline and psychological testing, coupled with an extensive post-injury rest period, is most likely to help properly identify and mitigate re-injury. Nonetheless, there is not yet professional consensus regarding the best standard diagnostic procedure or metric for TBIs or the professional qualifications required to make the decision that an athlete should be able to return to play. Should agreement on a best TBI diagnostic metric or clinical best practice emerge, legislatures may need to amend their statutes to adopt the proven practice, including the appropriate professional qualifications for those who provide clearance. Thus, Rhode Island's policy experimentation will provide useful data metrics, test a novel population-impacting approach to harm reduction, and clarify areas of scientific ambiguity.¹¹⁵

Another policy innovation adopted by two states imposes harsh penalties on coaches who fail to comply with the youth sports TBI law provisions. In Pennsylvania, the governing body of a school is required to suspend a coach for the remainder of a season if the coach fails to properly remove a student athlete from play.¹¹⁶ For a subsequent violation, a coach can have a longer suspension or even a "permanent suspension from coaching any athletic activity."¹¹⁷ Similarly, in Connecticut, the State Board of Education may revoke a coaching permit for a coach who fails to remove a student suspected of having a concussion from play.¹¹⁸ A coaching permit can also be revoked for a Connecticut coach's failure to complete training and refresher courses on concussion recognition and treatment.¹¹⁹ These types of provisions should help remove any sort of incentive that a coach might have to play a young athlete when the athlete has suffered a TBI, but few states have adopted these provisions. The reasons for most states' failure to consider such innovations are unclear.

Finally, states differ in the degree to which their youth sports TBI laws insulate various constituents from civil liability. Twenty-five jurisdictions have enacted youth sports TBI laws that attempt—directly or by inference—to limit liability for school districts, volunteers, healthcare providers, and others who might face lawsuits filed by athletes, their families, and others for damages on the basis of how these individuals respond to TBI-related events. These provisions usually have a carve-out for gross negligence and willful or wanton

115 This is precisely the type of policy innovation that, with respect to other statewide public health law matters, one might hope that the state policy laboratories would allow to flourish. However, as discussed earlier, the strong uniformity among state laws has made the "can states act as policy laboratories?" approach to lawmaking with respect to TBIs a rhetorical question.

116 24 PA. CONS. STAT. § 5323 (2013).

117 *Id.*

118 CONN. GEN. STAT. § 10-149c (2013).

119 *Id.*

misconduct.¹²⁰ A typical provision provides that a protected party “shall be immune from civil liability for good faith conduct arising from or pertaining to the injury or death of a student-athlete” if the party provides some proof that the conduct was in compliance with the law and “local school board policies relative to the management of concussions and head injuries.”¹²¹ Other statutes are more ambiguous, stating that the law does not create a “new cause of action” but not specifying whether other portions of state law might still provide for civil liability.

In contrast, the other twenty-five jurisdictions with youth sports TBI laws have determined (whether by silence or legislative debate) that the duties created by their youth sports TBI laws do not warrant special discussion of potential liability for key actors or immunity from civil liability. This decision to not provide waivers may have stemmed from the lack of evidence assessing whether liability waivers positively or negatively influence the identification and management of TBIs. This patchwork of liability waivers among types of parties and across states raises novel legal questions.¹²² For example, contests may involve athletes from different states or take place outside of one state’s jurisdiction under a different set of liability rules. In those circumstances, which state’s provisions will apply, and how do we know in advance?

F. Evidence, Influence, and Rushing: Common Causal Mechanisms in Youth Sports TBI Law Policymaking

While the specific motivation for youth sports TBI laws may be different in every state, there appear to be three common elements that motivated state lawmakers during the period studied here. Ideally, proposed legislation would be motivated solely by a broad set of public health law goals upon which all agreed, although realistically moderated by concerns such as fiscal constraint. However, for reasons explained earlier, the logic of youth sports TBI laws lacks a tight alignment with most scientific studies of the root causes and treatments of youth sports TBIs. This lack of alignment may be due to the actual motivating elements of the legislation, which have generally been threefold: the lobbying effort of the NFL, the sudden national attention given to the rise in youth concussions, and the desire to minimize the cost in implementing new state regulations covering a large portion of a state’s population and infrastructure. The combination of these

¹²⁰ See, e.g., IND. CODE § 20-34-7-5 (2013).

¹²¹ N.H. REV. STAT. ANN. § 200:51 (2013).

¹²² The connection between liability waivers and broad public policymaking on public health and other matters is a subject worthy of its own detailed analysis and cannot be fully explored here. However, it is certainly true that the relationship between waivers and policy efficacy has not been adequately studied, and one cannot be certain that such provisions will help lower the overall baseline TBI rates.

three elements provides one broad explanation for the rapid adoption of virtually identical laws across forty-nine jurisdictions. Each motivating element shall be examined in turn.

With respect to the NFL's role, public choice scholars have commented extensively on outside interest groups' roles in influencing state policymaking. Here, the states' legislative actions could be considered a form of regulatory capture.¹²³ In short, regulatory capture occurs when special interests compete to use government power for their own agenda.¹²⁴ Under the encompassing theory of public choice, government policymakers act for their constituents, as the "public," in creating an outcome, the "choice."¹²⁵ The outcome of a regulatory capture will represent the interest groups' agenda and (potentially) a form of government policymaking failure.¹²⁶ Obviously, this outcome is not always best for the public interest.¹²⁷

State-level lawmaking with respect to youth sports TBI laws suggests capture, though other alternatives are also plausible. In combination with the previously discussed absence of empirical evidence on which to base their policymaking, the states' legislative efforts could be captured—whether in rhetoric, strategic focus, or policy—by one powerful private interest group that was ready and willing to lead the charge. Therefore, if the policy "choice" has been captured, we should be concerned about whether this choice truly is the public will. That the NFL vigorously supported states' policymaking against an empirically ambiguous backdrop also raises concerns about how the resulting laws will operate and whether they will optimally benefit youth athletes. The laws may well prove to have a positive impact, and the NFL's role should then be commended. However, for reasons described earlier, the absence of evidence and data metrics suggests that it will be hard to determine in the short term what the impact has been.

The evidence gathered for this Article suggests that the NFL stepped into a policy vacuum with a compelling legislative message. This messaging built on the momentum generated by the events leading up to the Washington State Lystedt Law and ultimately shaped the policy debate from 2009 to present. To be fair, with any complex story, there are multiple causal explanations, and many of them deserve their own separate analysis. For example, it could be that states

123 See generally Sam Peltzman, *Toward a More General Theory of Regulation*, 19 J.L. & ECON. 211 (1976); Richard Posner, *Theories of Economic Regulation*, 5 BELL J. ECON. & MGMT. SCI. 335 (1974); George Stigler, *The Theory of Economic Regulation*, 2 BELL J. ECON. & MGMT. SCI. 3 (1971).

124 Michael E. Levine & Jennifer L. Forrence, *Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis*, 6 J.L. ECON. & ORG. 167 (1990).

125 *Id.* at 168.

126 *Id.* at 169.

127 *Id.* at 168–69.

tend to copy other states, and the youth sports TBI lawmaking process was no exception. Perhaps the uniformity and non-invasiveness of these laws is a reflection of consistent public opinion across states or a larger cultural norm in favor of more limited interventions into matters affecting sports and youth. In addition, fiscal concerns may have prevented a more aggressive policy intervention framework in most states. Or perhaps the NFL's "feel good" message provided motivation for legislators to align themselves with a low-negative, high-positive issue to add to their campaign portfolios. Any combination of these explanations could be true, and further research is needed to focus on them individually or collectively. For this Article's purposes, the policy uniformity and the restricted range of intervention are hypothesized to be the result of interest-group engagement and messaging about the problem. To that end, the evidence gathered here (and detailed in the Appendices) focuses on the role of a key interest group, the NFL, which had the largest ability to control the shaping of the message through media, legislative engagement, and public outreach.

One sign of the NFL's influence on the policy prescriptions of state youth sports TBI laws is the high prevalence of states that included the three tenets of model legislation in their youth sports TBI laws without diverging from the NFL's suggested form of intervention. In addition to the letters of support that NFL Commissioner Roger Goodell sent to state governors urging them to pass youth sports TBI laws similar to Washington's Zackery Lystedt Law,¹²⁸ the NFL also testified to legislative bodies in support of concussion safety measures.¹²⁹ Furthermore, in 2011, the NCAA announced that it would join the NFL's effort to encourage state legislatures to adopt legislation similar to the Lystedt Law.¹³⁰ Through these efforts, seven states passed some form of youth concussion legislation in 2010, twenty-seven states in 2011, nine states in 2012, and the remainder (absent Mississippi) in 2013.¹³¹

The evidence of the NFL's influence can best be seen by evaluating two broad metrics for determining the coupling between an interest group and the passage of laws aligned with the interest group's goals. To initially test for this influence, this Article notes the coupling of the interest group's name in news

¹²⁸ See *The Zackery Lystedt Law*, *supra* note 72.

¹²⁹ See, e.g., *NFL Urging States to Pass Youth Concussion Laws*, CBSSPORTS.COM, Feb. 23, 2011, <http://www.cbssports.com/nfl/story/14729129/nfl-urging-states-to-pass-youth-concussion-laws> ("NFL lobbyist Kenneth Edmonds urged DC council members to adopt rigorous concussion safety measures."). Details about the scope of testimony and other engagement can be found in the Appendix.

¹³⁰ Michael David Smith, *NCAA Joins NFL's Push for Youth Concussion Legislation*, NBC SPORTS: PRO FOOTBALL TALK, (June 7, 2011, 4:22 PM), <http://profootballtalk.nbcsports.com/2011/06/07/ncaa-joins-nfls-push-for-youth-concussion-legislation/>.

¹³¹ See *infra* Appendix 1.

accounts of the targeted legislation and/or its passage. In addition, the Article gathers data from a search of legislative history and/or public records surrounding the debate and eventual adoption of the state's youth sports TBI laws. In short, this initial examination of data can provide some assessment of the role of interest group influence in two critical realms—through news coverage of the youth sports TBI laws and through the legislative history of the enacted laws. Research in both realms conducted for this Article (and described in the Appendices) tracked and rated the strength of references to the NFL that appeared in conjunction with discussion of youth sports TBI laws.¹³² These two sets of indicators were compared to each other and analyzed according to the officially expressed support of the NFL for certain of the proposed laws.¹³³

As detailed in the Appendices, a high-strength news reference is defined by a direct connection between the state's youth sports TBI legislative history and the NFL's official lobbying, express support, recommendations, or sponsorship, as analyzed by the reporting news source.¹³⁴ High evidence of influence in the legislative history was established by one or more documents on the states' legislative websites with a direct connection between the state's legislative efforts and the NFL's official lobbying, express support, recommendations, or sponsorship.¹³⁵ In the context of legislative history, the term "direct connection" denotes that materials found in the legislative record were either provided or explicitly endorsed by the NFL and thereby reasonably could be regarded as evidence of the organization's lobbying for the bill.

References, direct and indirect, to the NFL were most common in the states with laws that received its express support.¹³⁶ The NFL was discussed in news accounts alongside the proposed youth sports TBI legislation in almost 75% of the states with such pending legislation.¹³⁷ All of the states with strong legislative history references to the NFL received its express support except for Washington State, where the Zackery Lystedt Law functioned as the foundation for subsequent legislation.¹³⁸ Even in states where the NFL did not expressly support the laws (as indicated in its public commentary, letters, and news references), the organization appears in the news and the legislative history.¹³⁹ The NFL is mentioned as a driving force for states' adoption of youth sports laws regardless

132 See *infra* Appendices 1–2.

133 See *Concussion Legislation*, *supra* note 77.

134 See *infra* Appendix 1.

135 See *infra* Appendix 2.

136 See *infra* Appendices 1–3.

137 See *infra* Appendix 1.

138 Information about specific references to the NFL in each state's legislative history is on file with the author and available upon request.

139 See *infra* Appendices 1–2.

of official influence in the development.¹⁴⁰

Clearly, the NFL made itself a visible lobbying force throughout states' legislative processes in developing and eventually passing youth sports TBI laws.¹⁴¹ The question is whether the organization's motivation in doing so diverged from the public interest. The issue of youth sports concussions is directly relevant to the NFL's private commercial goals in protecting the image and reputation of football. By lobbying and providing resources and recommendations for state legislatures, the NFL was able to frame the issue (reducing negative outcomes of concussions) and shape the response accordingly (facilitating overwhelmingly uniform laws and not directly regulating the content, rules, or procedures of football itself). One plausible reason for avoiding a direct focus on the frequency or severity of the concussions themselves might have been a large class-action lawsuit against the NFL that was pending throughout much of this period.¹⁴² That lawsuit focused on the dangers of both initial concussions and repeat concussions, while suggesting that the NFL's leadership failed to put measures in place designed to minimize the risk of initial and secondary concussions.

By their submission to the NFL's influence, the states failed to carefully calibrate policy for vulnerable constituents. Where the states should act either as careful fact finders or as labs for policy experimentation, the NFL's experience dealing with football concussions and the science surrounding incidences of TBIs in sports instead served as a form of reassurance to lawmakers who trusted the NFL's judgment about the connection between policy prescription and potential policy outcomes. Although some law may well be better than none, the states' actions in failing to broaden the scope of data upon which to legislate allowed the window for optimal experimentation and policy innovation to lapse.¹⁴³

Some suggest that the mere involvement of the NFL in meeting with state legislators, testifying in public hearings, and other such matters should not lead to a conclusion that its presence (direct and indirect) actually influenced legislative choices or policy outcomes. But there are two possible reasons why this is likely untrue. First, by defining the problem as "re-injury following re-entry", the NFL and its focus on the Lystedt story directed attention away from a different public health problem—the initial injury. Thus, focusing legislators on solving for the problem of injury after re-entry necessarily led away from more holistic policy

140 See *infra* Appendix 1.

141 See *infra* Appendix 1.

142 See, e.g., Andrew Brandt, *The Other Lawsuit*, MMQB WITH PETER KING (Nov. 9, 2013), <http://mmqb.si.com/2013/10/25/riddell-lawsuit-nfl-concussions-andrew-brandt/>.

143 The NFL is one of many important advocacy groups working to raise awareness and seek legislative resolution for the issues of youth concussions. The legislative influence of the NFL may also be consistent with the supporting efforts of many other organizations that also appear as advocates throughout the news stories and the states' legislative histories.

prescriptions designed to focus on initial prevention. Second, policymakers must be influenced by something, whether it is moral values, public opinion, culture, religion, or interest groups. And in the context of public health, direct personal contact with researchers and health advocates serves as the most important driver of policymaking choices in a given public health intervention.¹⁴⁴ Further, policy makers in public health settings tend to rely on sources that they trust and for whom they have access.¹⁴⁵ The NFL provided both access and a trustworthy message given its potential interest in reducing the overall incidence of concussions in sports generally, or at least football in particular. Thus, it seems more likely than not that the NFL's message and influencing served as one key motivation for the adoption of youth sports TBI laws.

The second motivation for the widespread adoption of youth sports TBI laws is the increased public understanding and awareness of the frequency of youth concussions. In addition to the concussion awareness campaign being led by the NFL,¹⁴⁶ other organizations have taken steps to raise awareness. For example, the National Academy of Neuropsychology and the National Athletic Trainers' Association partnered on a campaign to raise awareness on concussions and concussion treatment.¹⁴⁷ As a result of these efforts, more people—and certainly state legislators—became aware in recent years of the dangers that a concussion poses to a young athlete. Youth TBI laws protect children and raise awareness of an issue of public importance with few public opponents—factors that would typically reflect positively on a legislator's public persona and electoral positioning.

The third factor influencing the content and adoption of youth sports TBI laws is likely a desire to minimize state budget expenditures while providing a policy solution to a costly problem. This survey of youth sports TBI laws shows that the vast majority of states do not require new programs or costs to either state agencies or individual school districts.¹⁴⁸ Rather, the concussion policies are

144 See, e.g., Simon Innvaer et al., *Health Policy-Makers' Perceptions of Their Use of Evidence: A Systematic Review*, 7 J. HEALTH SERVICES RES. & POL'Y 239, 239–44 (2002) (reviewing studies and interviews with health policymakers and identifying personal contact as a key policy-choice making variable).

145 See, e.g., Christopher A. Jewell & Lisa A. Bero, *Developing Good Taste in Evidence: Facilitators of and Hindrances to Evidence-Informed Health Policymaking in State Government*, 86 MILBANK Q. 177 (2008).

146 See generally Roger Goodell Q&A: "We Can Make the Game Safer", NFLREVOLUTION.COM (Sept. 5, 2012), <http://www.nflrevolution.com/article/Roger-Goodell-Q-038-A-8216-We-can-make-the-game-safer-8217-?ref=1299>.

147 National Academy of Neuropsychology (NAN) and National Athletic Trainers' Association (NATA) Team up on Campaign to Raise Concussion Awareness, NHL.COM (Sept. 9, 2009, 5:03 PM), <http://www.nhl.com/ice/news.htm?id=499020>.

148 See, e.g., COLO. LEGIS. COUNCIL STAFF, STATE AND LOCAL REVISED FISCAL IMPACT, S.B. 11-040, (Colo. 2011), available at <http://www.leg.state.co.us/clics/clics2011a/csl.nsf/fsbillcont3/>

expected to increase a state agency or school's workload but not require additional funding.¹⁴⁹ In fact, the NFL on its *Frequently Asked Questions about Concussion Prevention Laws* webpage¹⁵⁰ explains that the cost of implementing a concussion prevention and awareness bill similar to the Zackery Lystedt Law is zero since the three tenets contain "no requirements that resources be spent to hire or train medical professionals or to purchase equipment."¹⁵¹ Additionally, general information on effective concussion management practices is publicly available and free through the Centers for Disease Control for high school and youth coaches, parents, athletes, and school professionals.¹⁵²

Through the combination of these three motivating factors, all of which relate to the NFL's influence, states rapidly adopted largely identical youth sports TBI laws. While enactment of a youth sports TBI law modeled on the Lystedt law is generally a positive step toward reducing the number of traumatic brain *re-injuries* among young athletes, these laws could still be improved in a variety of ways—particularly by focusing on primary prevention. It is troubling that states only began to act rapidly on this issue by adopting NFL-endorsed policies from a consistent one-size-fits-all model once the NFL's organizational lobbying began. This kind of interest-based TBI-prevention policy choice framing could result in solutions that primarily reflect the interest groups' concerns rather than the public interest overall. In addition to the legislative innovations above, there are a variety of other ways that state legislative interventions could more effectively address the problem of youth concussions. Section IV now turns to those interventions.

IV. FAILURES OF EXISTING POLICY AND POTENTIAL POLICY IMPROVEMENTS

The aforementioned interventionist public health law approaches to concussions have many common features, a few innovations, and a high degree of uniformity. All but a few suffer from three critical shortcomings. First, they fail to include evaluative metrics to determine whether the law's reforms are helping to solve the problem. Second, existing youth sports TBI laws have a singular focus on reducing the secondary effects of concussions rather than attacking their root causes. Finally, existing youth sports TBI laws fail to track individual athletes and the rise in risk associated with athletes who suffer multiple concussions. These factors may undermine the efficacy of existing

A9CE9CEE12645CAA8725780800800D80?Open&file=SB040_r1.pdf.

¹⁴⁹ *Id.*

¹⁵⁰ *Frequently Asked Questions About Concussion Prevention Laws*, NFLHEALTHANDSAFETY.COM, <http://nfl-www.fkhstaging.com/map/FAQ.php> (last visited Dec. 5, 2013).

¹⁵¹ *Id.*

¹⁵² *Id.*

youth sports TBI laws; solutions are proposed here to remedy that effectiveness gap and potentially address root causes. These are just some of an array of potential policy innovations, but their efficacy can only fully be valued by policymakers with a more robust and consistent research agenda and use of data—which at the moment all but a few states do not collect or analyze for trends.¹⁵³

A. *Mandatory Aggregate Reporting and Feedback Regimes*

The most perplexing omission in youth sports TBI laws is the failure of almost all states to develop a reporting and testing system to evaluate the effectiveness of the laws. One would think that testing, evaluation, and science should be at the core of any broad-based public health initiative. But only three states have (indirectly) mandated a surveillance program by implementing regulation or mandatory general reports.¹⁵⁴ These provisions or regulations will allow forward-thinking states to test and report on the effectiveness of their newly created youth concussion regimes, while others will not have a systematic method of evaluating the law's impact. While this would entail a cost to the entity charged with producing the report, an analysis of the youth concussion data is the primary way to determine whether the law is actually effective. Since states have recently reduced their general expenditures drastically,¹⁵⁵ state legislators may be relying on the press to investigate the effectiveness of the youth concussion measures in lieu of a state study, though there is not specific evidence to support this contention.¹⁵⁶ Given that one purpose of state-based concussion lawmaking could have been to help reduce both the incidence and long-term health consequences of concussions, this failure to include an evaluative metric will reduce the amount of information available to

153 See, e.g., Abby S. Haynes et al., *Galvanizers, Guides, Champions, and Shields: The Many Ways that Policymakers Use Public Health Researchers*, 89 MILBANK Q. 564 (2011).

154 See, e.g., MO. REV. STAT. § 167.775 (2013) (“Any statewide athletic organization with a public school district as a member shall be required to publish an annual report relating to the impact of concussions and head injuries on student athletes which details efforts that may be made to minimize damages from injuries sustained by students participating in school sports.”).

155 See *Fiscal Survey of States: Spring 2011*, NAT’L GOVERNORS ASS’N & NAT’L ASS’N OF STATE BUDGET OFFICERS, at viii (2011), http://www.nasbo.org/sites/default/files/Spring-2011-Fiscal-Survey_0.pdf (“Twenty-three states made budget cuts to their fiscal 2011 budgets totaling \$7.8 billion. Thirty-nine states made mid-year budget cuts of \$18.3 billion in fiscal 2010, while 43 states made \$31.3 billion in mid-year cuts in fiscal 2009.”).

156 However, an innovative grant-funded program was recently developed to begin making such an assessment of effectiveness. See *Evaluating Implementation of Return-to-Play Laws for Athletes with Concussions to Increase Effectiveness of Existing and Future Legislation*, ROBERT WOOD JOHNSON FOUND., <http://www.rwjf.org/en/grants/grant-records/2011/11/evaluating-implementation-of-return-to-play-laws-for-athletes-wi.html> (last visited Dec. 5, 2013).

policymakers to determine whether and how the law is helping to reduce the incidence of primary and secondary TBIs.¹⁵⁷

B. Mandatory Individual Tracking Metrics

Existing youth sports TBI legal regimes fail to include player-level tracking systems designed to ensure that a youth athlete's medical history is tightly tied to treatment and outcomes.¹⁵⁸ Because the epidemiology of TBIs is relatively consistent with respect to the increased level of risk for athletes who have suffered repeated concussions, states have a strong interest in monitoring the overall number of concussions sustained by athletes. First, with respect to short-term health, medical professionals generally agree that the recovery time for concussions increases when one has suffered a previous concussion. Second, as described earlier, CTE and other long-term health outcomes appear to be tied to a history of repeated concussions. Only by tracking the concussion history for each student athlete can a state's public health officials, school districts, coaches, and parents be certain that the medical response to an athlete's concussion is appropriately tied to the athlete's previous history. The CDC agrees, identifying "collecting data from schools" and "studying changes in concussion knowledge . . . before and after the policy is put in place."¹⁵⁹ Vermont and Massachusetts are pioneers in this effort, but other states have not taken this initiative. It is true that such efforts could be costly, but given the already required intervention for return-to-play clearance, it hardly seems more onerous to notate a school's master file or record about the occurrence. But assuming that a state's true concern is cost, others have suggested that the CDC develop and implement a national surveillance system, thus potentially removing cost as a barrier.¹⁶⁰

C. Required Baselines

Professional leagues and some states also rely on pre/post-concussion evaluative baseline player metrics to determine when a player's cognitive

157 The failure to include such information in reporting could be explained as a cost-saving mechanism. In addition, the visibility of such data could complicate policy adherence and overall enforcement, since individual schools or coaches might face additional scrutiny if such rates were substantially higher than the state's average.

158 However, Massachusetts and Michigan youth sports TBI laws suggest that such a function is possible under their regimes. MASS. GEN. LAWS ch. 111, § 222(b) (2013); MICH. COMP. LAWS §§ 333.9155, 333.9156 (2013).

159 See Nat'l Ctr. for Injury Prevention & Control, *Get a Heads Up on Concussion in Sports Policies: Information for Parents, Coaches, and School & Sports Policies: Information for Parents, Coaches, and School & Sports Professionals*, CDC, <http://www.cdc.gov/concussion/pdf/HeadsUpOnConcussionInSportsPolicies-a.pdf> (last visited Dec. 5, 2013).

160 See, e.g., *Sports-Related Concussions*, *supra* note 47.

functions substantially deviate from a pre-established “norm”; players demonstrating such deviations would not be permitted to reenter athletic competition until their cognitive functioning returned to normal. In professional sports, baseline tests are the norm for players.¹⁶¹ Across youth sports, such tests are not yet common practice. The evaluation of TBIs at the professional level is useful as a guiding lesson for high school sports programs but may also be a financially unrealistic model for states absent partnerships with universities, donors, and others who provide the baseline testing vehicles.¹⁶² Yet some states such as New Jersey and Rhode Island have still managed to develop regimes that build in baseline testing. However, as discussed earlier, the effectiveness of baseline testing is not empirically certain.

D. Direct Intervention in Particular Sports

All but one or two provisions across the entire set of state youth sports TBI laws are targeted to reducing *secondary* TBI risks—those short- and long-term health consequences that begin after an athlete has already suffered a concussion. Reading pamphlets about how to identify when players have a concussion, making them sit out when a concussion is suspected, and having them evaluated by a doctor are all well-timed and important state-law reforms, but they all take the existence of a concussion for granted.

Why this near-exclusive focus on what happens *after* a TBI has already occurred? Part of the failure to address the underlying issue itself—prevention of initial TBIs—is that to do more would require a more invasive legislative and financial mandate than some states may be willing to implement.¹⁶³ An additional factor is rooted in the nature of competition itself. The sports that are most likely to cause concussions (cheerleading, football, hockey, and soccer) are all popular team sports cutting across regional, gender, and cultural lines. Therefore, while it is obvious that not offering these sports reduces the likelihood of concussions within a given youth population, such an option might be politically and culturally untenable.¹⁶⁴ The focus of TBI laws on results of concussions, instead

161 See, e.g., *Standardized Concussion Tests Coming to the NFL*, NFL, <http://www.nfl.com/news/story/09000d5d81e76bed/article/standardized-concussion-tests-coming-to-the-nfl> (last updated July 26, 2012).

162 New Jersey’s approach to baseline testing is unique. It applies a small surcharge to car-registration fees, which is then used, in part, to pay for baseline neuropsychological testing of high-school student athletes. See N.J. STAT. ANN. § 18A:40-41.7 (West 2013).

163 Because states were so focused on the NFL-endorsed three tenets, scant evidence exists that any substantial alternatives to the Lystedt model were considered. That having been said, a few states, discussed herein, did engage in some policy innovation.

164 This being said, President Theodore Roosevelt once vigorously and successfully campaigned for more rigorous safety standards in football, spurring substantive game-changing rules (like the invention of the “forward pass”), as well as the creation of national sport oversight

of reducing concussions to begin with, may also have to do with the social value of youth sports activities and desire to avoid directly changing the customary rules of a game.¹⁶⁵ That youth sports TBI laws primarily focus on what happens *after* a TBI presupposes the legitimacy of the event that caused the TBI. That sports, and particularly football, have nationally recognized social (and perhaps even political) import and value speaks again to why a powerful private interest group like the NFL would seek to influence the debate and shape the solution from the beginning.

Aside from the extreme position of eliminating popular sports, there are still other direct sport interventions that likely would reduce the overall number of concussive events, including the mandatory improvement of equipment and changes in style of play or specific in-sport maneuvers.¹⁶⁶ Yet while the exact science of how helmets reduce or minimize risk of concussions is still uncertain, it is probable that newer helmets will help eliminate that risk more effectively than older ones.¹⁶⁷ A legislative mandate that organized football programs buy entirely new helmets every two to three years may reduce the number of concussions in youth football, but requiring new helmets so often would also substantially drive up the costs to schools and school districts.¹⁶⁸ Because helmets are an expensive part of the game, teams have incentives to minimize their constant replacement. However, because helmets can be partially re-manufactured for a much lower cost than a new helmet (both new materials

regimes like the NCAA. Jim Morrison, *The Early History of Football's Forward Pass*, SMITHSONIAN.COM (Dec. 28, 2010), <http://www.smithsonianmag.com/history-archaeology/The-Early-History-of-Footballs-Forward-Pass.html>.

165 For example, researchers recently used advanced MRI techniques to show that frequent heading of the ball in soccer resulted in TBI-like symptoms. See, e.g., *Frequent Soccer Ball "Heading" May Lead to Brain Injury*, ALBERT EINSTEIN COLL. MED. OF YESHIVA UNIV. (June 11, 2013), <http://www.einstein.yu.edu/news/releases/915/frequent-soccer-ball-heading-may-lead-to-brain-injury/>. However, because heading the ball is a core component of soccer play enjoyed by fans and players alike, it is difficult to imagine a law-based regime aimed at changing that sport's core play not suffering from public resistance.

166 The NFL and other sports leagues have, in recent years, focused on penalizing certain head-impacting maneuvers, such as when the NFL began penalizing players in 2013 who engaged in running or tackling by leading with the crown of their helmets. See, e.g., Ian Rapoport, *Rule Change Banning Head-On Approach Comes with Questions*, NFL, <http://www.nfl.com/news/story/0ap1000000151942/article/rule-change-banning-headon-approach-comes-with-questions> (last updated Mar. 20, 2013).

167 See, e.g., Jenny Vrentas, *The First Line of Defense*, MMQB WITH PETER KING (Oct. 22, 2013), <http://mmqb.si.com/2013/10/22/nfl-helmets-head-injury-concussion/> (describing safety features implemented in newer helmets and the current competitive market for helmet innovation and safety).

168 Although there is not a direct correlation between age of helmet and TBI prevention, once helmets reach a certain age, effectiveness as even a baseline head injury protection decreases. See, e.g., Schwarz, *supra* note 63 (examining helmet standards in youth sports and noting that "[m]ore than 100,000 children are wearing helmets too old to provide adequate protection").

and/or new design), older helmet designs can remain in circulation for far longer than some experts believe is useful.¹⁶⁹

A full discussion of the “helmet debate” is beyond the scope of this Article. Nonetheless, a growing body of literature suggests the lack of long-term evidence regarding the quality and effectiveness of football helmets in preventing TBIs.¹⁷⁰ As those studies explain, football helmet standards have not changed meaningfully since 1973,¹⁷¹ and the studies undertaken to test their effectiveness do not approximate real world conditions. Further, helmet standards are regulated by a private consortium and are not subject to rigorous reporting requirements. Therefore, it is not certain that existing helmets are particularly safe. In this highly uncertain environment, when one considers that mandatory helmet replacement may also be particularly difficult for low-income parents and athletes, one can see how the status quo is maintained.

Finally, with respect to sports-based interventions, youth sports TBI laws have generally taken a one-size-fits-all approach. They do not focus on how changes in play in individual sports might reduce the primary instances of youth sports TBIs. There may be many reasons for this,¹⁷² but the laws do not acknowledge the scientific consensus that youth TBIs vary on the basis of age, the type of sport, and whether the athlete is male or female.¹⁷³ In the future, there may be value in legislation that addresses organized sports risks in a more specific manner. For example, because a significant proportion of TBIs suffered by female youth student-athletes occur in soccer and cheerleading, more narrowly or finely tailored laws might direct state officials to involve soccer or cheerleading sport-specific associations in the development of TBI identification and treatment guidelines. Or, because most male youth student-athlete TBIs occur in soccer and football, TBI laws may maximize their impact by heightening sport-specific TBI reduction techniques. It is also possible that race and socioeconomic factors play a role, resulting in a race- or class-skewed level of TBI incidences across sports. However, lacking data beyond anecdotes, it is

169 For example, in an incident described by Lazarus, an NFL player’s helmet design was originally developed in 1988, but was being worn many years later when an incident occurred, despite subsequent improvements to helmet design. See Arthur Lazarus, *NFL Concussions and Common Sense: A Recipe for Medical Errors and a Lesson for Physician Leaders*, 37 PHYSICIAN EXECUTIVE J. 6 (2011) (describing an NFL game where two players with on-field concussions were removed and then re-inserted in the same game, despite both of them having typical concussive indicators).

170 McCrory et al., *supra* note 94, at 255 (suggesting no proven link to preventing TBIs).

171 See, e.g., Alan Schwarz, *Soon, Helmet Data at a Keystroke*, N.Y. TIMES, Jan. 9, 2011, <http://query.nytimes.com/gst/fullpage.html?res=9805E2D9143DF93AA35752C0A9679D8B63>.

172 See Jewell & Bero, *supra* note 145.

173 Andrew E. Lincoln, *Trends in Concussion Incidence in High School Sports: A Prospective 11-Year Study*, 39 AM. J. SPORTS MED. 958 (2011).

difficult to determine whether this is broadly true.

Separately, legislative mandates about how the games must be played (for example, eliminating certain types of hits or maneuvers, or reducing the amount and/or intensity of practice drills for certain sports) would also certainly reduce risk and drive overall concussion rates downward. The CDC agrees, identifying “rule changes” and banning or limiting “certain drills or techniques” in sports in an attempt to reduce injury.¹⁷⁴ However, such mandates would directly interfere with the inner workings of a sport and would be vigorously challenged by those who play such sports. These three direct intervention alternatives (eliminating certain sports, imposing additional equipment expenses, changing rules) can each reduce the overall number of concussions, but not without potential public backlash or limitations due to fiscal constraints across districts and states. Therefore, it has been easier for lawmakers to focus on secondary interventions, where the core nature of the sports themselves is not directly threatened. Evidence presented here should lead states and districts to experiment with an array of approaches, which would allow policymakers to gather evidence to determine which interventions, changes, and regulations are the most effective at reducing incidences both initial and repeat youth sports TBIs.

E. Fixing the Reporting Disincentive

Most analysts agree that the incentives to clear concussed players to return to play places teams, doctors, and athletes in asymmetrical positions.¹⁷⁵ Sound long-term medical opinions cannot always be provided in a high-pressure situation where quick decisions are needed. As such, the more that state-level reforms can successfully disentangle these incentives, the more successful these reforms are likely to be in reducing short-term re-injury and longer-term consequences. However, when a physician or health professional has a management position with a team, the tension between protecting patients (players) and maximizing team value (winning the game) seems like an obvious conflict that should be avoided. Certainly, coaches face similar incentives to potentially prioritize winning over player safety.¹⁷⁶ As noted above, two states have attempted to address this concern by creating punitive enforcement mechanisms for failing to report TBIs. Another way might be to create an anonymous reporting structure, modeled after whistleblower provisions in other laws. The provision could allow parents, athletes, and other officials to share information when they have reason to believe that the state’s TBI mandates are not being followed.

174 See Nat’l Ctr. for Injury Prevention & Control, *supra* note 159.

175 See, e.g., Michael McCrea et al., *Unreported Concussion in High School Football Players: Implications for Prevention*, 14 CLINICAL J. SPORT MED. 13 (2004).

176 See, e.g., Brendan L. Smith, *Benched Coaches Shooting Back*, A.B.A. J. (Oct. 1, 2010, 2:00 AM), http://www.abajournal.com/magazine/article/benched_coaches_shooting_back/.

F. Learning from Professional Sports and Independent Associations

There is no shortage of conclusions, speculations, and recommendations about how to assess the TBI problem, minimize its effects, and prevent injuries in both adults and children. We now know that concussions occur in youth sports more frequently when inexperienced athletes are playing in team sports.¹⁷⁷ We also know that ice hockey and football participants are the most likely to suffer from concussive events.¹⁷⁸ And we know that professional sports organizations have undertaken tough new measures designed to minimize the risk and consequences of initial concussions, and also to prevent multiple TBIs. These measures include changes in equipment, sports rules, or the times and locations in which sports are played.¹⁷⁹ The lessons of these varied approaches can be instructive.

Both the NFL¹⁸⁰ and the National Hockey League (NHL) have changed rules with the explicit intent of reducing TBIs.¹⁸¹ At the collegiate level, the NCAA has partnered with the CDC to promote concussion safety and best practices.¹⁸² Major League Baseball's aggressive concussion protocols require: (1) baseline testing for all players and umpires, (2) strict evaluation of individuals who have suffered concussions, and (3) the creation of a seven-day disabled list for players with concussions.¹⁸³ Independent associations also have stricter post-concussive guidelines than are implemented by state laws. For example, the American Academy of Pediatrics (AAP) cautions that a youth athlete who has suffered a concussion should be gradually reintroduced into athletic activities over a five-day period to ensure that there are no residual signs of concussive effects.¹⁸⁴ In youth sports organizational bodies, similar efforts are underway to establish clear guidelines and recommendations for various sports.¹⁸⁵ Most recently, Pop Warner

177 Mitka, *supra* note 18, at 1775–76.

178 2009 *Participation – Alphabetically*, NAT'L SPORTING GOODS ASS'N, http://www.nsga.org/files/public/2009_Participation-Alphabetically_4Web_100521.pdf (last visited Dec. 6, 2013).

179 Christine Provvidenza & Charles H. Tator, *Sports Injury Prevention: General Principles*, in CATASTROPHIC INJURIES IN SPORTS AND RECREATION: CAUSES AND PREVENTION—A CANADIAN STUDY 59 (Charles H. Tator ed., 2008).

180 See *Memo Explains Policy to Have Trainers Monitor for Concussions*, NFL, <http://www.nfl.com/news/story/09000d5d82547e65/article/memo-explains-policy-to-have-trainers-monitor-for-concussions> (last updated July 26, 2012).

181 See Lazarus, *supra* note 169.

182 *Attention College Sports Fans: CDC and NCAA Team Up on Concussion Safety*, CDC, http://www.cdc.gov/concussion/sports/cdc_ncaa.html (last updated Mar. 9, 2012).

183 See *Erick Almonte Slept Under Supervision*, ESPN, <http://sports.espn.go.com/mlb/news/story?id=6437072> (last updated Apr. 27, 2011).

184 See Mark E. Halstead, *Sport-Related Concussion in Children and Adolescents*, 126 PEDIATRICS 597, 604–05 (2010).

185 See, e.g., AACCA *Concussion Management and Return to Play Protocol*, AM. ASS'N OF

(an umbrella organization for youth football, cheer, and dance programs in forty-two states) created new standards to limit contact during football practice and heighten concussion awareness among participants.¹⁸⁶

In comparison to the approaches adopted by professional sports and other national associations, states have adopted minimalist strategies, particularly with respect to return-to-play provisions. Recall that the science is fairly clear even in its ambiguity: there is simply no strong evidence to determine when an athlete is well enough to return to competition or which intervention method following a TBI works best to speed recovery.¹⁸⁷ And while the MLB and AAP's timetables are considered scientifically reasonable, the scientific evidence about the scope and length of post-concussive re-entry is still disputed.¹⁸⁸ At a minimum, one might suggest that the graduated assessment regimes of professional sports organizations should be a floor, not a ceiling, for state law innovations involving vulnerable athletes. Yet no state law sets out a graduated timetable for return to play or evaluation, instead relying on the now-standard minimum 24-hour waiting period. This is not to say that relying on professional sports standards is a policy panacea. For example, return-to-play decisions in the NFL are not uniformly followed and could actually serve as negative guidance or modeling for youth programs.¹⁸⁹ Nonetheless, the policies created by professional and independent associations should not be entirely ignored by states.

G. *Evaluating Policy Alternatives*

Each of the alternative policy reforms described above addresses substantive lawmaking choices in existing youth sports TBI laws, suggesting that more can be done by way of innovating with TBI reduction measures in youth sports TBI laws.¹⁹⁰ Based on the evidence presented here, a truly comprehensive youth sports TBI prevention framework should reduce the overall number of youth head injuries, prevent secondary head injuries, and potentially reduce the (as yet unknown) long-term effects of head injuries caused by athletic competition. But

CHEERLEADING COACHES & ADMINISTRATORS, <http://www.aacca.org/content.aspx?item=/Resources/concussions.xml> (last visited Dec. 5, 2013).

186 See *Pop Warner Leagues Limit Contact in Practice*, SI.COM (June 12, 2012, 8:41 PM), <http://sportsillustrated.cnn.com/2012/more/06/12/pop.warner.contact.ap/index.html>.

187 Mitka, *supra* note 18, at 1775–76.

188 Recall that one expert testified that brain-injury proteins linger up to three months after initial injury. See notes 64–66 and accompanying text.

189 See Lazarus, *supra* note 169.

190 This discussion focuses only on reduction of short-term health consequences and non-fatal TBI-related events. Morbidity and mortality are commonly associated with all types of TBIs in children, yet reforms in youth sports TBI laws do not directly address the study of morbidity and mortality with respect to youth TBIs developed outside of athletics and compare interventions for those TBIs with ones that might be appropriate in youth sports. See, e.g., Carol A. Hawley, *Behaviour and School Performance After Brain Injury*, 18 BRAIN INJURY 645, 645–49 (2004).

the real issue is the overall scope of reduction in youth TBIs, not whether some negligible rate reduction can be achieved through new state laws. Nonetheless, there are those who argue that alternative frameworks—or no frameworks at all—would be better than the laws studied here.

Prior to the enactment of the Zackery Lystedt Law in Washington State, public health interventions and interest group efforts, while well-meaning, did not appear to have a measurable effect on the overall volume of concussions in youth populations.¹⁹¹ During the era preceding the Zackery Lystedt Law, the CDC set forth vigorous, clear, and helpful guidelines that were available to coaches, youth athletes, and their families. Yet without the drive, knowledge, and enforcement power of state law to ensure that this information reached those populations, there is little evidence that any of the information had a substantive effect on the goal of reducing TBIs in youth populations.

Such motivation, information, and power have recently been provided, in part by the NFL, and seem to have finally encouraged the states to act. In this post-Zackery Lystedt Law era of public health law reform, the issue has been dominated by the NFL's effective interest group advocacy. The regulatory capture of youth sports TBI lawmaking by such a powerful interest group is one key reason why the underlying policy discussion has failed to exhibit any real experimentation, and why the resulting policies lack the innovative spark that states' independent legislative debates could have provided.

CONCLUSION

We know that youth TBIs and their residual effects are a serious problem, with a complex web of causal factors and ensuing consequences that remain poorly understood. Legislative mandates have been enacted in all but a single state with the express purpose of providing information and establishing minimum responses to TBI events, presumably with the laudable goal of reducing the overall number of youth TBIs, minimizing their residual and repeated impact, and providing for safer outcomes in youth sports. All of these goals are worthy of public praise, and some state action was surely better than state silence. All agree that youth sports place youth at risk of suffering TBIs, yet public health laws fail to foster assessment and reporting mechanisms to evaluate and minimize these well-known risks and health outcomes. The Zackery Lystedt Law and the three tenets it helped to define provided a useful framework to *begin* analyzing how the law might help address youth sports TBIs, but it does not have to be the end goal.

As the above analysis shows, the three tenets are simply not enough to

191 It is possible that, despite the documented rise in reported TBIs, such efforts did reduce the pace of that rise. Unfortunately, there are no accurate measures to test this hypothesis.

accomplish the long-term youth TBI reduction goals that states could be (and arguably should be) attempting to reach. It may be that heightened attention to these secondary prevention measures could reach back to some initial prevention solutions, but without knowing precisely what those initial prevention measures are, it remains unclear how state law could successfully reach toward them. Laws that require a mandatory sitting-out period, awareness by coaches, parents, and players of concussion signs, and medical evaluation of concussed players all represent worthy initiatives that focus on helping recognize when an athlete is already in a concussive or post-concussive state. These initiatives do not appear on their face to address TBIs in the first instance—i.e., the reduction of the risk factors that might cause the initial concussion itself. Because legislative focus on the Lystedt Law's three tenets has achieved relatively uniform post-concussive initiatives, it may be time to consider whether this unique state/private-actor partnership should be expanded to include additional interest groups, broader metrics, and additional and more comprehensive proposals for reform. It seems clear that the NFL's influential role at least partly explains why more innovation has not flourished, despite the differences in politics, culture, and sports across the 49 states and the District of Columbia.¹⁹² This is not to say that these laws accomplished nothing, but to suggest that the definition of the problem, helped along by the NFL, has not led to substantive discussion about policy alternatives or a broader framing or focus on youth sports TBI's more generally.

Further study of youth sports TBI laws over time is needed to test their efficacy and both secondary and primary risk reduction. But to do so most effectively, either more policy innovation or more evaluative metrics are required at the state level. Ideally, research would take a state-by-state survey of the field, examine variation in various state youth sports TBI laws, and somehow evaluate the successful or innovative traits across such laws by comparing longitudinal policy outcomes in the populations targeted by these laws. However, as has been explained, existing laws exhibit very little variance across states. Therefore, such research might take the form of case studies, whereby researchers would select a few states with radically different youth sports TBI laws, with another as a control, and research how states with stronger, mid-range, weaker, or no laws protected or enhanced public health outcomes with respect to youth sports TBIs. Additional research could take a quantitative approach, drawing upon data trends across states (such as rates of concussions pre/post-enactment) and could determine whether the regimes have had some effect.

Unfortunately, comprehensive evaluative and innovative public health law

192 In other contexts, this phenomenon could be considered a form of groupthink, causing policy paralysis in state youth TBI lawmaking as a result of a lack of dissent or competing voices in the debate. *See, e.g.,* IRVING L. JANIS, *VICTIMS OF GROUPTHINK: A PSYCHOLOGICAL STUDY OF FOREIGN-POLICY DECISIONS AND FIASCOES* (1972).

studies of youth sports TBI laws and their outcomes are made extraordinarily difficult by the failure of states to include robust reporting mechanisms in their youth sports TBI laws. If policymakers are serious about using the force of the law to have an impact on public health, they must also create evaluative metrics to ensure that their lawmaking has the desired effect on public health outcomes. The lack of required evidence gathering for such an important policy goal is both surprising and counterintuitive. If the stated goal of such legislation is to reduce the overall number of TBIs and to minimize their long-term consequences, the only way that one can test whether a state's law is effective is by (a) creating a baseline assessment for all student athletes, (b) creating a reporting regime that requires coaches, hospitals, and schools to publicly identify instances of youth TBIs across a variety of variables (sports, school districts, etc.), and/or (c) repeating baseline assessments on student athletes over time to measure any potential decline in performance and determine whether such decline might be caused by an athlete's injury. The long-term effort to reduce youth sports TBIs must involve policy evaluation, policy experimentation, or measurement of policy outcomes in youth populations. Only when states choose to gather and assess information tied to their policies can they best determine whether the policies are helping to achieve their goals—and if not, how they might change the policies to be more effective. Even though all states but one now have such youth sports TBI laws, the policymaking game is not over. The risk-reducing steps taken in response to the public health concerns described here were just the beginning of a long season that is far from over.

APPENDICES¹⁹³*Appendix 1: Evidence of Influence—News*

A LexisNexis news search was run twice, once from major newspapers and once from regional news sources across all fifty states and the District of Columbia. The search included a terms & connectors search across the last five years for “National Football League” or NFL in the same paragraph as brain or concussion or TBI, student or youth or child, and legis! or bill or act.¹⁹⁴

Out of the then 45 states (including the District of Columbia) with youth sports TBI laws, 33 were mentioned in combination with references to the NFL in these news sources. These references were assigned strengths of high, medium, or low.

A low-strength reference includes only a passing mention of the name of the state in a list while discussing ongoing legislative efforts across other states, albeit in combination with some mention of the NFL in the greater discussion.¹⁹⁵ Seven states have an NFL news reference strength of low.¹⁹⁶

A medium-strength reference includes mention of the state in a statement connecting it with the NFL but without an explicit reference to a lobbying, supporting, or otherwise sponsoring relationship. Sometimes this involves mention of a former NFL player’s involvement in the legislative efforts.¹⁹⁷ Other times this involves naming the state in a list, similar to the low-strength references, but qualifying the list of states by suggesting that they have been influenced by the NFL in an accompanying remark.¹⁹⁸ Nine states have an NFL

193 This analysis was conducted in January and February of 2013, before several states passed youth sports TBI laws or changes to existing TBI laws.

194 (“National Football League” or NFL) /p (brain or concussion or tbi) /p (student or youth or child!) /p (legis! or bill or act) and date geq (12/11/2007)

195 For example: “Other football-laden states like Alabama[,] . . . Nebraska and Oklahoma have all passed youth-concussion legislation within the past two years.” Shannon J. Owens, *Florida Lagging Behind on Youth-Concussion Prevention*, ORLANDO SENTINEL, Aug. 1, 2011, http://articles.orlandosentinel.com/2011-08-01/sports/os-shannonowens-youth-concussions-20110801_1_concussions-youth-sports-league-rich-football-history.

196 The seven states are Alabama, Iowa, Minnesota, Nebraska, Oklahoma, and Wyoming, as well as the District of Columbia.

197 For example: “Becker, a former offensive lineman and nine-year NFL veteran, was in Springfield [Illinois] Wednesday to promote legislation . . .” Jeff Engelhardt, *Former Bear Raises Concussion Awareness*, CHI. DAILY HERALD, Feb. 23, 2011, <http://www.dailyherald.com/article/20110223/news/702239919/>.

198 For example: “[L]awmakers in . . . Connecticut, Massachusetts . . . and Rhode Island legislatures are debating similar laws. . . . All of these states have or are contemplating adopting some version of the National Football League rules” Juanita Thornton, *Brain Injuries Are Serious*, ROANOKE TIMES, Mar. 2, 2010, at A13.

news reference strength of medium.¹⁹⁹

A high-strength reference includes a direct connection between the state's legislative efforts and the NFL's official lobbying, express support, recommendations, or sponsorship.²⁰⁰ Seventeen states have an NFL news reference strength of high.²⁰¹

Twelve states had no NFL news references. Although these states were sometimes referred to in published news articles, the referenced article was unrelated to the accompanying NFL term that prompted a "match" under the search protocol. Sometimes this occurred when one state's newspaper discussed other states' legislative activities in relation to the NFL. Other times this happened when an article combined two topics, discussing both youth sports TBI laws (and referencing the NFL in combination with such) as well as related traumatic brain injury resources available in other states.²⁰²

Appendix 2: Evidence of Influence—Legislative history

A legislative history search was conducted through each of the 45 state legislatures' websites.²⁰³ Although the websites are not uniform in their layout or content, the laws can all be found through a few methods. Some states' sites offer an embedded search with preset query boxes for the document type (such as bill, senate or house file, or legislative document), number, and year. Other states' sites offer clickable links for legislative history with searchable indexes organized thereunder. Still others states' sites offer only a general search query box. Most states' sites offered two or more of these search methods.

Once the enrolled law was found as a document on the legislative websites, the preceding or linked pages usually provided accompanying documents. The availability and depth of these resources varied greatly. Law-related documents included sponsor statements, statements of intent, fiscal notes, minutes and audio, committee actions, sectional analyses, suggested amendments, news stories, and academic articles, among others. Where available, these documents were

199 The nine states are Connecticut, Illinois, Massachusetts, New Mexico, Oregon, Rhode Island, South Dakota, Utah, and Virginia.

200 For example: "Adding support to the legislation at the Capitol press conference were Joe Browne, senior adviser to NFL Commissioner Roger Goodell" *Stevenson Joins Effort to Reduce Youth Sports Injuries*, PA. HOUSE REPUBLICAN CAUCUS (Jan. 28, 2011), <http://www.pahousegop.com/NewsItem.aspx?NewsID=10371>.

201 The 17 states are Arkansas, Arizona, California, Colorado, Florida, Kentucky, Louisiana, Mississippi, Missouri, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Texas, Washington, and Wisconsin.

202 The twelve states are Arkansas, Delaware, Hawaii, Idaho, Indiana, Kansas, Maine, Maryland, Nevada, North Carolina, North Dakota, and Vermont.

203 As described above, *see supra* note 193, this analysis was conducted in early 2013, prior to several states' adopting or changing existing youth sports TBI laws.

searched for any mention of NFL or National Football League.

Out of then 45 states (including the District of Columbia) with youth sports TBI laws, 14 had references to the NFL in these legislative history sources.²⁰⁴ These references were assigned strengths of high, medium, or low, describing the evidence of influence that the NFL had on the legislative process.

High evidence of influence in the legislative process is shown by one or more documents with a direct connection between the state's legislative efforts and the NFL's official lobbying, express support, recommendations, or sponsorship. The term "direct connection" denotes that materials found in the legislative record were either explicitly endorsed or provided to the state legislature by the NFL and reasonably could be considered as evidence of the organization's apparent lobbying for the bill. For example, in Alaska, high evidence of influence included attached documents in the legislative history for Alaska Statutes section 14.30.142, such as a "Youth Concussion Education Pack," distributed to the state legislature and expressly endorsed by the NFL, as well as minutes from the House Education Committee that show the NFL Director Kenneth Edmonds reading a prepared statement into testimony.²⁰⁵ Such direct lobbying efforts appeared in ten states.²⁰⁶

Medium evidence of influence in the legislative history is shown by references to the NFL in documents that suggest the organization's influence but do not confirm direct involvement in the efforts to pass state laws. The term "suggest" denotes materials found in the legislative record that did explicitly mention the NFL but that were neither clearly printed nor provided to the state legislature by the organization itself. Such materials were assessed as having the potential to influence the legislation without concluding that the NFL played a direct role in lobbying. For example, within the bill analysis and fiscal impact statement for Florida Statutes section 943.0348, a summary history of the Zackery Lystedt law mentions that "Roger Goodell, Commissioner of the National Football League, sent a letter to state governors urging their support of legislation that would better protect young athletes by mandating a more formal and aggressive approach to treatment of concussions."²⁰⁷ This reference connects the NFL to Florida's legislative efforts but does not clearly indicate either that the letter was sent to Florida's governor or that this legislation was begun on

204 Information about specific references to the NFL in each state's legislative history is on file with the author and available upon request.

205 Student Athlete Traumatic Brain Injuries: Hearing on H.B. 15 Before the H. Educ. Standing Comm., 27th Leg., 1st Sess. 11 (Alaska 2011) (statement of Kenneth Edmonds, Director, Nat'l Football League).

206 These states are Alaska, California, Delaware, Hawaii, Kansas, Nebraska, Nevada, North Dakota, and Texas, as well as the District of Columbia.

207 PROF'L STAFF OF THE BUDGET COMM., BILL ANALYSIS AND FISCAL IMPACT STATEMENT, S.B. 256 (Fla. 2012).

account of such letter. Similar evidence appeared in three states.²⁰⁸

Low evidence of influence in the legislative history is shown by a passing and indirect mention of the NFL's connection to the law in the documents available with the bill online. For example, Maryland's legislative record has one mention of the NFL embedded in a summary of background for the bill. Within the fiscal and policy notes is a reference to "the online ImPACT test, which is the same test used by the National Football League."²⁰⁹ Maryland was the only state with such low evidence.

The failure of the vast majority of states to upload supporting documentation and information pertaining to their youth sports TBI laws is worth noting. Whether this lack of policymaking source material is intentional or accidental, it undoubtedly frustrates governmental transparency, as well as this and future attempts at online legislative history research. Thirty-one states have youth sports TBI laws for which the legislative history online reveals no NFL references.²¹⁰ However, lack of evidence is not dispositive of legislative history NFL references for the majority of states, given the absence of any additional information about the legislative process pertaining to those states' youth sports TBI laws.

Appendix 3: NFL Statements Regarding Its Own Legislative Involvement

The NFL describes its influence in the development of these youth sports TBI laws on its website, which lists legislation by state.²¹¹ The site was "last updated August 16, 2012," but curiously specifies that "[a]s of October 2012, 40 states (plus the District of Columbia and the city of Chicago) have adopted youth concussion laws."²¹² The page goes on to add that "[t]he NFL supports and recognizes the laws as they represent the main principles of the Lystedt Law model legislation..." and lists the three tenets.²¹³ The rest of the page lists "Legislative updates on all 50 states."²¹⁴ Confusingly, the list includes 44 youth sports TBI laws passed.²¹⁵

From this list, the NFL expresses its support for 34 of the bills, and the states' bill summary bears the marker "Status: Legislation passed with NFL

208 These states are Florida, Michigan, and Washington.

209 See *supra* note 204.

210 These states are Alabama, Arkansas, Arizona, Colorado, Connecticut, Iowa, Idaho, Illinois, Indiana, Kentucky, Louisiana, Maine, Massachusetts, Minnesota, Missouri, North Carolina, New Hampshire, New Jersey, New Mexico, New York, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Dakota, Utah, Virginia, Vermont, Wisconsin, and Wyoming.

211 See *Concussion Legislation*, *supra* note 77.

212 *Id.*

213 *Id.*

214 *Id.*

215 *Id.*

support,” with a few minor wording differences among them.²¹⁶

For two states, the NFL specifies that the legislation passed with support from affiliated parties, but not from the organization directly: Illinois, “with support from the Chicago Bears,” and Washington, “with Seattle Seahawks’ support.”²¹⁷

For eight states, the NFL does not indicate their support in any direct manner.²¹⁸ However, the NFL does include a summary of the legislation, specifying that each includes or contains the “three tenets of model legislation.”²¹⁹

For Wyoming alone, the NFL does not indicate its support, and furthermore specifies how this youth sports TBI law diverges from the three tenets of the model legislation.²²⁰ This is consistent with the NFL’s uniform promotional approach via the distribution of materials and representatives supporting the model legislation.²²¹ The specification of differences between the Wyoming bill and the model legislation could be read to suggest the NFL’s disapproval.

The express support of the NFL corresponds extremely well with those states with the strongest news and legislative history NFL references.²²² Of the 14 states with any legislative history NFL reference at all, all of them except Washington received express NFL support.²²³ Of the 17 states with high-strength news NFL references, all of them except Washington received express NFL support.²²⁴ This is consistent with the express lobbying efforts of the NFL in relation to these states’ bills and the open news coverage of this influence. The Washington exception is easily explained by virtue of its status as the original state from which the model legislation was drawn.

The modified support of the NFL through an affiliated party corresponds with states where there was either medium- (Illinois) or high- (Washington) strength news NFL references. Washington also had medium strength legislative history NFL references.

The absence of expressed NFL support corresponds somewhat with those

216 *Id.* These states are Alaska, Alabama, Arkansas, California, Colorado, Delaware, Florida, Hawaii, Iowa, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maryland, Maine, Mississippi, Minnesota, Missouri, North Carolina, North Dakota, Nebraska, New Hampshire, New Jersey, Nevada, New York, Ohio, Pennsylvania, South Dakota, Texas, Utah, Vermont, and Wisconsin, as well as the District of Columbia.

217 *Id.*

218 *Id.*

219 *Id.* These states are Connecticut, Massachusetts, New Mexico, Oklahoma, Oregon, Rhode Island, and Virginia.

220 *Id.*

221 *See supra* Appendix 2.

222 *See supra* Appendices 1–2.

223 *See supra* Appendix 2.

224 *Id.*

states having the lowest news and legislative history NFL references.²²⁵ Six of the eight states receiving no statement of NFL support have medium-strength news references.²²⁶ The other two states, Oklahoma and Wyoming, had low-strength news references.²²⁷ Maryland was the only state with a low-strength legislative history NFL reference and where the NFL expressed support of the state's bill.²²⁸ Similarly, of the seven states with low news NFL references, only Oklahoma and Wyoming's bills are not expressly supported by the NFL.²²⁹ Of those 13 states with no news NFL references at all, the NFL page does not recognize Arkansas as having pending or passed relevant legislation at the time of the update, while all the rest received the NFL's express support.²³⁰

In summary, where the legislative history shows any evidence of the NFL's influence, this is consistent with the NFL's expressed support. Likewise, where the news reports show strong evidence of the NFL's influence, this is consistent with the organization's expressed support. The NFL's lack of expressed support is consistent with those states with low- or medium-strength news references, but not all states with no, low-, or medium-strength news references lacked express NFL support. This is understandable given the limitations of both the press to cover legislative history in depth and the search protocol applied to return all relevant news references.

Ultimately, to whatever extent the NFL has promoted youth sports TBI legislation, it has overwhelmingly succeeded. In states where there are youth sports TBI laws, the NFL has supported a vast majority of them, and it has played an influential role in the legislative history of many. The NFL may have had a similarly strong influence over additional states' legislative efforts, but the records available for research online are either nonexistent or fail to comprehensively disclose this role.

225 *Id.*

226 *See supra* Appendix 1.

227 *Id.*

228 *Id.*

229 *Id.*

230 *Id.*

Human Research Subjects As Human Research Workers

Holly Fernandez Lynch*

ABSTRACT:

Biomedical research involving human subjects has traditionally been treated as a unique endeavor, presenting special risks and demanding special protections. But in several ways, the regulatory scheme governing human subjects research is counter-intuitively less protective than the labor and employment laws applicable to many workers. This Article relies on analogical and legal reasoning to demonstrate that this should not be the case; in a number of ways, human research subjects ought to be fundamentally recast as human research workers. Like other workers protected under worklaw, biomedical research subjects often have interests that diverge from those in positions of control but little bargaining power for change. Bearing these important similarities in mind, the question becomes whether there is any good reason to treat subjects and protected workers differently as a matter of law. With regard to unrestricted payment, eligibility for a minimum wage, compensation for injury, and rights to engage in concerted activity, the answer is no and human subjects regulations ought to be revised accordingly.

* Executive Director, Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School. J.D./M.Bioethics, University of Pennsylvania. My thanks to Glenn Cohen, Einer Elhauge, Carl Elliott, Jonathan Jay, Sophia Lee, Paul Lombardo, Kristin Madison, Michelle Meyer, Collin O'Neill, Ben Roin, Benjamin Sachs (HLS), Benjamin Sachs (NYU), Seema Shah, Jeff Skopek, and Patrick Taylor for very helpful comments and discussion. Thanks also to participants in the 2012 Harvard Law School Health Law Policy Workshop; the Petrie-Flom Center 2012 Conference on The Future of Human Subjects Research Regulation; and the American Society of Law, Medicine, and Ethics 2012 Health Law Professors Conference. Finally, thank you to Jordan Bachr and Afroditi Giovanopoulou, who provided research assistance, and the excellent editors of the Yale Journal of Health Policy, Law, and Ethics.

TABLE OF CONTENTS

INTRODUCTION	125
I. LEGAL AND ETHICAL REQUIREMENTS FOR HSR	130
II. RESEARCH PARTICIPATION AS A JOB.....	132
A. THE BIOETHICISTS	132
1. HEALTHY SUBJECTS AND AVOIDING EXPLOITATION	133
2. EQUATING SUBJECT PAYMENT WITH A WAGE.....	137
3. ANALYSIS	139
B. LEGAL BACKGROUND.....	142
1. EMPLOYEES AND EMPLOYERS.....	142
2. THE EMPLOYMENT STATUS OF HUMAN RESEARCH SUBJECTS.....	144
3. ANALYSIS	148
III. CLARIFYING THE ANALOGY AND ITS SCOPE	149
A. WHY NOT APPLY WORKLAW DIRECTLY TO SUBJECTS?	149
B. SHARED CONCERNS OF SUBJECTS AND PROTECTED WORKERS	151
1. DIVERGENT INTERESTS	151
2. BARGAINING POWER, COLLECTIVE ACTION, AND EXPLOITATION.....	152
3. DEMOCRATIC DEFICITS, DEPENDENCY, AND THE NEED FOR PROTECTION	153
C. WEAKER ANALOGY, SAME PROTECTIONS	154
IV. LEVELING SUBJECTS UP TO WORKERS	156
A. PAYMENT.....	156
1. PAYMENT CEILING	159
2. MINIMUM WAGE	163
3. UNEMPLOYMENT.....	165
B. CARE AND COMPENSATION FOR INJURY	166
1. SCOPE, NEED, AND BURDEN.....	170
2. COVERED HARM AND CAUSATION.....	173
3. FUNDING, TORT PREEMPTION, AND ADDITIONAL CONCERNS.....	175

4. UNPAID SUBJECTS	177
C. WORKING CONDITIONS AND INSPECTIONS	179
D. COLLECTIVE BARGAINING AND UNIONS.....	181
V. OUTSTANDING OBJECTIONS	191
CONCLUSION	193

INTRODUCTION

In the early 1930s, a physician named William Osler Abbott was working to develop a new technique for rapidly intubating the human intestine. He swallowed the tubes himself until he “was sick of the sight of them” and worked out some aspects of the technique using hospital patients. Eventually, however, he realized the need to experiment on healthy human subjects.¹

Abbott planned to offer two dollars a day to anyone willing to take the job, and expected no difficulty recruiting at the height of the Depression. However, relief administrators refused to refer anyone for this type of work. Abbott lamented that “[i]n that day of poverty, destitution and collapsing homes, not one solitary subject did I ever get through the very agencies that were apparently striving to find jobs, to make jobs, to beg jobs of any kind for hungry men.” Abbott also failed to recruit “beggars,” “tramps, vagrants, and the unemployed,” all of whom recoiled at the work once it was described to them. Eventually, he enlisted the help of a “black janitor on the hospital floor,” offering a finder’s fee for every healthy subject brought in “appearing sober” and fasting, and finally had some success.²

As he prepared to exhibit the intubation technique at a medical conference, however, Abbott faced an unexpected problem: his subjects went on strike, seeking more money. At the eleventh hour, he managed to recruit new volunteers from a class of medical students, “a shipment of scab labor . . . that would have made any factory foreman green with envy.” Then, “the National Labor Relations Board being as yet unborn,” Abbott “had the pleasure of indulging in a little old-fashioned capitalism. We fired the whole lot of [strikers], lock, stock and barrel. The exhibit went off like clockwork.”³ Thereafter, Abbott began advertising in newspapers, which generated enough interest that he was able to have his pick from among the applicants, most of whom volunteered for financial reasons.⁴

Abbott’s encounters are in many ways from a different time, but much of what he describes could apply equally well to biomedical research involving human subjects today: advertisements, payment for participation, and recruitment based on physical characteristics and willingness to cooperate. One key feature, however, stands in stark contrast. Abbott clearly viewed his subjects as workers and their research participation as a type of job. But this view of the subject has never taken hold. Today, researchers, ethicists, and regulators generally see subjects as a class unto themselves – a class in need of special protection as a

1 W. Osler Abbott, *The Problem of the Professional Guinea Pig*, 68 TRANSACTIONS AM. CLINICAL & CLIMATOLOGICAL ASS’N 1, 1 (1957).

2 *Id.* at 2-3.

3 *Id.* at 4.

4 *Id.* at 6-8.

result of a history of research scandal and other supposedly unique characteristics, such as the nature of subjects' vulnerability, motivation, and risk taking.

As a result, the protections offered to many traditional workers under U.S. labor and employment law (hereinafter, "worklaw") have been treated as completely separate from and exclusive of the regulatory scheme governing clinical research. And although that regulatory scheme, and the thrust of research ethics in general, is heavily focused on subject protection, some traditional workers – usually those satisfying the legal definition of "employee" – receive a number of legal protections that research subjects do not.

In recent years, however, the idea of extending worker-type legal protections to human subjects has received some attention from bioethicists, and a few legal decisions have even treated subjects as employees. In fact, while the subject strike remains a novelty, the threat has been successfully leveraged at least once since Abbott's time by healthy subjects in a drug metabolism study in Philadelphia.⁵

These developments raise a number of interesting questions: In what ways are human subjects really like other workers? How are workers more protected than subjects, and when do the rationales for various worklaws fit the context of human subjects research (HSR)? Should human subjects be treated like employees, independent contractors, or volunteers, or is a more nuanced application of worker protections needed, if at all? Each of these questions is addressed in what follows, but a number of prefatory statements are in order.

First, although the overarching discussion may be relevant to the entire realm of HSR including social and behavioral research, primarily due to space constraints, the focus here will be limited to biomedical research – and more specifically, to clinical trials conducted in competent adult subjects.

In that context, this Article seeks to critically examine the way HSR is currently regulated and to challenge the status quo approach that treats subjects and workers as wholly distinct entities. Unlike much existing scholarship,⁶ however, it does not proceed by first identifying a variety of problems with the HSR regulations and then searching for analogous models from clinical medicine or elsewhere that potentially could be combined to do better; this is not an Article about the inadequacies of institutional review boards, the problems associated with poorly worded consent forms, or the inherent shortcomings of the HSR regime. In fact, in this Article, the question is not about better versus worse, adequate versus inadequate, good versus bad, but rather consistency versus inconsistency. Thus, the Article starts with the premise that subjects and protected workers are incredibly similar, and then uses that fundamental similarity to identify problems with the existing HSR regulatory structure –

5 Robert P. Helms, *Human Guinea Pigs "Band Together" and Win a Pay Hike*, GUINEA PIG ZERO (2003), <http://www.guineapigzero.com/helmsstrike.html> (last visited Nov. 21, 2013).

6 See *infra* Part II.A.

aspects that are problematic by virtue of their inconsistency with the legal treatment of protected workers, and that could be rendered less problematic through a consistent approach. Thus, the scope of the question includes only subjects and traditional workers, not subjects and every other conceivable comparator.⁷

Using this particular type of analogical methodology, the Article advances a simple argument: (1) Research subjects are, in relevant respects, like those workers entitled to the protections offered by worklaw. In the clearest cases, they are healthy, paid, and treat research participation as work. But even beyond that, with the exception of true altruists, *all* types of clinical research subjects share the key features that motivate worker protections, including interests that diverge from those in positions of control, poor bargaining power, and collective action problems. (2) Thus, the two groups should be treated alike unless there is some good reason to behave otherwise. At present, however, they are treated inconsistently and in both directions – subjects are simultaneously more and less protected than traditional workers (note that I did not say “better” and “worse,” as that is a distinct normative inquiry). The former scenario is largely tabled for future work in order to permit sufficient attention to those circumstances in which subjects currently find themselves *less* protected than their counterparts in the working world. This is the more appropriate starting point given its contrast against the intense drive for subject protection that motivates both the existing legal and predominant ethical paradigms.

Of course, establishing a strong analogy between subjects and protected workers is only one step in the analysis; analogy alone cannot resolve the essential question of how to address inconsistent treatment between two similar points of comparison. Thus, it is also necessary to determine which side of the analogy should be destabilized in order to establish consistency – should we aim to level subjects up to workers, granting them new baseline protections that run alongside those they already have, or workers down to subjects, eliminating some of their current baseline entitlements?

To that point, there is substantial scholarly and social disagreement as to the value and importance of the minimum wage, workers’ compensation, and union activities, for example, and certainly some abuses. Opponents of these worklaw protections maintain that they are blunt instruments for nuanced work

⁷ The worker comparison is chosen here primarily because it has been raised on numerous occasions in the bioethics literature, but has received insufficient legal attention. *See, e.g.*, Paul McNeill, *Paying People to Participate in Research: Why Not?*, 11 *BIOETHICS* 390 (1997); Benjamin Sachs, *The Exceptional Ethics of the Investigator-Subject Relationship*, 35 *J. MED. & PHIL.* 64 (2010); Marx W. Wartofsky, *On Doing It for Money*, in NAT’L COMM. FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH, APPENDIX TO REPORT AND RECOMMENDATIONS: RESEARCH INVOLVING PRISONERS (1976); and *infra*, Section II.A. However, selecting this analogy is not meant to foreclose others, such as comparisons to patients, investors, donors (of body parts, time, or money), surrogates, prostitutes, prisoners, soldiers, or even lab rats.

relationships and the variety of workplaces covered and that they unjustifiably hinder the economy, such that we should instead leave everything to contract, relegating the terms and conditions of work to the market. Similarly, some have suggested a private ordering approach to certain aspects of HSR,⁸ allowing competent adult subjects to contract with researchers and sponsors for the benefits and protections they need to make research participation worthwhile.

However, for workers and subjects alike, there are reasons to be wary of a purely market-based approach. As we will see, both lack information available to the “opposing” party, both face substantial constraints on their bargaining power, and both are susceptible to other problems recognized by behavioral economists, such as weighting present interests more strongly than future interests and otherwise failing to act in full accordance with the “rational actor.” These features also help to explain why, if subjects wanted or needed additional protections, they have not already attained them.⁹

Recognizing that this is a point of contention, this Article assumes (and I believe) that we ought to be extremely hesitant to scale back the hard-fought protections offered by worklaw and won over time by the labor movement. Leveling workers down is inadvisable, whereas leveling subjects up is consistent with the overall protective approach to HSR. To clarify the thesis, then, subjects and workers are like cases *and subjects ought to be treated more like protected workers*, unless the reasons for granting a particular type of worker protection fail to apply to research subjects. And because this Article tables the issue of whether subjects should retain those protections that are more extensive than those available to traditional workers, the focus here is on changes that would be additive to the status quo research regulations.

In setting forth this argument, the Article makes a number of unique contributions. It is the first to offer a comprehensive legal analysis of this analogy and the first to dig deeply into the question of why we protect workers the way we do in order to determine whether those reasons are similarly applicable to the subjects of biomedical research. It also offers new support for the argument against research exceptionalism (i.e., the idea that research is special), and sets a foundation for future applications of the worker analogy to determine the protections that ought to be extended to organ donors, surrogate parents, and the like.

With these broader goals in mind, the Article proceeds as follows: Part I provides brief background on the current regulations governing HSR, and identifies several reasons why these regulations evolved independent of worklaw. Part II then describes and critiques existing bioethical and legal analyses of the application of worker protections to subjects. Although these threads in the

8 See, e.g., Michelle N. Meyer, *Regulating the Production of Knowledge: Research Risk-Benefit Analysis and the Heterogeneity Problem*, 65 ADMIN. L. REV. 237 (2013).

9 Another important possibility is simply that the model of subjects as unique has become so entrenched that subjects themselves have not generally thought to look to worklaw as a point of comparison.

literature provide an important starting point, they fail to adequately acknowledge that workers are not a homogenous group, nor are they all entitled to legal protection. Thus, Part II endeavors to fill that gap by considering the different types of workers recognized by the law, the standards for distinguishing between them, and the reasons those distinctions matter. It concludes that at least some types of subjects, namely those who are healthy and paid to participate in nontherapeutic research, can almost certainly be legal employees – the most protected category – without changing a thing; this conclusion itself is an important and novel contribution to the literature, since it is decidedly not how subjects are treated at present.

However, Part III explains that simply treating subjects as employees would be less than ideal, as it would incorporate existing problems from worklaw into the research setting. Most importantly, the employee label currently fails to capture all workers in need of legal protection, and would similarly fail to include all subjects in need of worker protection. Thus, rather than analogizing to a flawed status quo, Part III explores the fundamental reasons that certain types of workers *ought* to be legally protected, many of which are sources of vulnerabilities shared by a variety of research subjects. Finally, this Part argues that even those subjects who do not closely resemble traditionally protected workers (i.e., those subjects who volunteer altruistically, for therapeutic benefits, or without pay) should nonetheless receive worker protections either because they cannot be reliably distinguished from other subjects or because they face the same relevant vulnerabilities.

From the starting point that workers should not be stripped of important protections, Part IV considers the remaining options of leveling subjects up or identifying relevant differences between the two groups. With regard to payment, it concludes that the current implicit limits on subject compensation are inappropriate and should be lifted, and that some subjects should be offered at least a modified minimum wage but not unemployment compensation. It argues that injured subjects should be guaranteed a no-fault remedy along the lines of the worker compensation system, with some adjustments regarding the types of injuries that will qualify. In contrast to regulation of the workplace, research site conditions are best handled by application and enforcement of existing building codes, facilities regulations, and the like. And although subjects are likely to face some difficulty joining together for collective bargaining and alternative approaches are likely to be more fruitful, they should be protected in their concerted activities and any attempts at unionization.

Finally, Part V briefly responds to a few outstanding objections, emphasizing the crucial point that protecting subjects like workers is not meant to encourage “professional guinea-pigging,” but rather to ensure that like cases are treated alike. In the end, human research subjects ought to be fundamentally recast as human research workers.

I. LEGAL AND ETHICAL REQUIREMENTS FOR HSR

There are many different types of human subjects research, and studies may be risky or less so, biomedical or behavioral, early or later phase, government-funded or private. There are also many different types of human subjects, who may be paid or unpaid, healthy or sick, vulnerable or less so. Nonetheless, with few exceptions,¹⁰ the regulations governing HSR do not differentiate on these grounds. Instead, the regulations generally cast a wide and undifferentiated net, protecting the many types of human subjects in the many types of research studies in two primary ways.¹¹

First, they require that researchers obtain and document subjects' consent to research participation under conditions that minimize the possibility of coercion or undue influence and after a number of specific disclosures.¹² Second, in order for potential subjects to even be given the opportunity to participate, an Institutional Review Board (IRB) comprised of disinterested scientific and nonscientific members must first review and approve proposed research. More specifically, IRBs cannot allow research to proceed unless risks to subjects will be minimized and remaining risks are reasonable in relation to anticipated benefits to subjects or society; subject selection will be equitable; informed consent will be sought; data will be adequately monitored to ensure subject safety; subject privacy and confidentiality will be protected; and there will be sufficient protections for vulnerable groups.¹³ IRBs are the ultimate gatekeepers in the world of HSR, authorized to approve, require changes to, reject, suspend, or terminate research studies.¹⁴

This heavily protectionist approach to HSR is best understood in historical context as a response to a number of ethical scandals and human rights tragedies that have been well documented elsewhere,¹⁵ such as the Nazi experiments, Tuskegee syphilis study, and others. This helps to explain, at least in part, why the HSR regulatory scheme has evolved separately from worklaw as

10 Note that there are some specific regulatory requirements for research involving pregnant women, children, and prisoners. See, e.g., 45 C.F.R. pt. 46, subparts B, C, D (2013); 21 C.F.R. pt. 50, subpart D (2013).

11 See 45 C.F.R. pt. 46 (the U.S. Department of Health and Human Services' version of the "Common Rule"); 21 C.F.R. §§ 50, 56, 312, 601, 812 (the relevant FDA regulations). The Common Rule covers most federally funded HSR due to its adoption by a number of government agencies. See *Federal Policy for the Protection of Human Subjects*, U.S. DEP'T. OF HEALTH & HUM. SERVICES (HHS), <http://www.hhs.gov/ohrp/humansubjects/commonrule/index.html> (last visited Nov. 21, 2013). Investigation of a product regulated by FDA will be subject to FDA's (very similar) regulations, potentially in addition to the Common Rule.

12 45 C.F.R. §§ 46.116-.117; 21 C.F.R. §§ 50.20, 50.25, 50.27.

13 45 C.F.R. § 46.111; 21 C.F.R. § 56.111. In addition to initial review of research, IRBs must also engage in continuing review at least annually. 45 C.F.R. § 46.109(e); 21 C.F.R. § 56.109(f).

14 45 C.F.R. §§ 46.109(a), 46.113; 21 C.F.R. §§ 56.109(a), 56.113.

15 See ADVISORY COMM. ON HUMAN RADIATION EXPERIMENTS (ACHRE), FINAL REPORT 98-104 (1995). Part I of the ACHRE report also includes a discussion of the historical evolution of the HSR regulatory regime.

a non-work regime: the Nazi victims and Tuskegee subjects hardly resembled workers in the usual sense, but rather looked like victims of torture and patients who had been subject to exploitation and neglect. Of course victims and workers should be protected differently.

Another factor leading to this segregated approach is that the HSR regulations were built on a predominantly medical model. Physicians were the ones who “took the lead in drawing up rules of conduct for human subjects research, and they were concerned to make the rules consistent with medical ethics and comparatively unconcerned to create rules consistent with the rules we apply in nonmedical contexts.”¹⁶ As a result, investigators – like practicing physicians – have traditionally been expected to act protectively toward their subjects’ interests as a matter of professional responsibility,¹⁷ in contrast to the often antagonistic relationship between management and labor. Moreover, the clinical context in which biomedical research takes place also seems to create an intuition that research participation cannot be work, such that worklaw approaches would be inapposite.¹⁸

In reality, doctors engaged in research have fundamentally different goals and responsibilities than doctors engaged in medical practice, but this distinction has taken a long time to take root and remains somewhat controversial.¹⁹ And even if one accepts that distinction, there is no consensus that research participation actually constitutes work, or that it should be treated as such.²⁰ Ultimately, neither the victim model nor the clinical paradigm leaves much room for a labor approach to HSR, from either legal or clinical perspectives.

As noted above, one obvious result of the fact that participation as a human subject traditionally has been treated as outside the realm of worklaw is that subjects are protected in a number of ways that workers, and others engaged in

16 Sachs, *supra* note 7, at 76.

17 See, e.g., *Moral Science: Protecting Participants in Human Subjects Research*, PRESIDENTIAL COMM’N FOR THE STUDY OF BIOETHICAL ISSUES (PCSBI) 9, 70-74 (Dec. 2011), <http://bioethics.gov/sites/default/files/Moral%20Science%20June%202012.pdf> [hereinafter *Moral Science*] (recommending that human research protections be more explicitly understood as professional standards).

18 On the other hand, the subjects of prison research, also of critical interest in the 1960s and ‘70s when the foundation for the current regulations was laid, may have lent themselves more favorably to a work-based regulatory regime, considering that they were often healthy, paid for their participation, involved in commercial research, and in many cases signed waivers acknowledging research risks. See Valerie H. Bonham & Jonathan D. Moreno, *Research with Captive Populations: Prisoners, Students, and Soldiers*, in THE OXFORD TEXTBOOK OF CLINICAL RESEARCH ETHICS 461, 465 (Ezekiel J. Emanuel et al. eds., 2008). However, the unique nature of their institutional confinement and restrictions on free choice weakened any resemblance to traditional work.

19 See Howard Brody & Franklin G. Miller, *The Clinician-Investigator: Unavoidable but Manageable Tension*, 13 KENNEDY INST. ETHICS J. 329 (2003); Sachs, *supra* note 7, at 76-77.

20 See McNeill, *supra* note 7, at 391.

risky activities, are not.²¹ Indeed, we do not have IRBs to approve which lawful employment opportunities we may be offered, nor do the agencies responsible for occupational health and safety and avoidance of employment discrimination consider various factors that are part of IRBs' mandate, such as social value, nonphysical risks, fair distribution of burdens and benefits, and the like. Moreover, in most other realms of decision-making, we speak only of consent, which may often be implicit, not of explicit, written "informed consent."²² There are a host of disclosures that must be made to employees regarding their rights and certain types of workplace hazards, but these pale in comparison to the requirements for research consent.²³ These and other examples of research exceptionalism skewed toward heavy subject protection have received quite a bit of attention in the literature, alongside more and less convincing arguments regarding potential justifications, as well as calls for reform.²⁴

The flip side, however, has received quite a bit less attention: to the extent that the HSR regulations take a protectionist approach, and to the extent that research subjects share important features with traditional workers, should the protections extended to subjects really be *less* stringent than those provided by worklaw? In many cases, the answer is no.

II. RESEARCH PARTICIPATION AS A JOB

*A. The Bioethicists*²⁵

Although far from a movement, a few bioethicists have considered the protections offered by worklaw as they search for ways to improve flaws they have identified in the existing human subjects regulatory regime.²⁶ One line of argument notes the potential benefits of a labor framework given the current state

21 Again, note that suggesting research subjects are more protected than others does not necessarily correspond to a normative judgment as to whether they are better protected.

22 Franklin G. Miller & Alan Wertheimer, *Preface to a Theory of Consent Transactions: Beyond Valid Consent*, in *THE ETHICS OF CONSENT: THEORY AND PRACTICE* 80 (Franklin G. Miller & Alan Wertheimer eds., 2010).

23 See, e.g., 29 C.F.R. §§ 500.75-.76 (2013) (describing disclosures that must be made to migrant agricultural and seasonal day-haul workers); 29 C.F.R. § 1910.1200 (2013) (OSHA's Hazard Communication Standard). Some have argued for expanding the use of mandatory disclosure in employment law. See, e.g., Cynthia Estlund, *Just the Facts: The Case for Workplace Transparency*, 63 *STAN. L. REV.* 351 (2011).

24 For example, Ben Sachs considers the ethical principles applicable to research subjects that are not applicable to employees and volunteers, but not the legal differences per se. Sachs, *supra* note 7. See also Miller & Wertheimer, *supra* note 22, at 79-105; James Wilson & David Hunter, *Research Exceptionalism*, 10 *AM. J. BIOETHICS* 45, 48-49 (2010) (and the related open peer commentaries in the *American Journal of Bioethics*, Vol. 10, No. 8 (Aug. 2010)).

25 Note that not all of the authors cited in this section would necessarily self-identify as bioethicists. However, that title is used as shorthand to refer to arguments that have been made outside of the legal sphere.

26 Ari VanderWalde & Seth Kurzban, *Paying Human Subjects in Research: Where Are We, How Did We Get Here, and Now What?*, 39 *J.L. MED. & ETHICS* 543, 553-55 (2011).

of paid participation by healthy subjects, while another emphasizes the discrepancy between the treatment of subjects and other wage-earners, and considers the implications of a unified approach, without necessarily endorsing it. However, a number of critical questions remain.

1. Healthy Subjects and Avoiding Exploitation

One reason to adopt a labor approach for HSR is the simple fact that as a result of payment, at least some subjects view research participation as a job,²⁷ and even those who do not may sometimes reasonably be viewed as selling their bodily services for money, just like other laborers.

In addition to reimbursement for out-of-pocket costs, subjects of all types may be given additional payment, characterized as compensation for time and inconvenience, an enrollment incentive, or even just a token of appreciation.²⁸ However, whereas patient-subjects in a potentially therapeutic trial spanning the course of two years might receive between \$25-50 per monthly visit, healthy subjects in a phase I drug study lasting a few weeks might receive \$200-400 per day.²⁹

The difference can be explained by the fact that for healthy subjects in nontherapeutic research, the risks and burdens of participation are not compensated by the prospect of direct medical benefit or even contribution to an area of personal relevance. Thus, substantial payments are often essential to adequate recruitment.³⁰ However, it is worth noting that while research participation as a healthy subject may be unpleasant, it is not usually as risky as the general public perceives it to be.³¹ Subjects are closely monitored and serious adverse drug reactions are rare. Nonetheless, they do occur, certain types of

27 ROBERTO ABADIE, *THE PROFESSIONAL GUINEA PIG: BIG PHARMA AND THE RISKY WORLD OF HUMAN SUBJECTS RESEARCH* 2, 78 (2010); *GUINEA PIG ZERO: AN ANTHOLOGY OF THE JOURNAL FOR HUMAN RESEARCH SUBJECTS* 5-6 (Robert Helms ed., 2002); Matthew Weinstein, *A Public Culture For Guinea Pigs: US Human Research Subjects After the Tuskegee Study*, 10 *SCIENCE AS CULTURE* 195, 204 (2001); Nancy Ondrusek, *Making Participation Work: A Grounded Theory Describing Participation in Phase I Drug Trials from the Perspective of a Healthy Subject* 78-80, 178 (2010) (unpublished Ph.D. dissertation, University of Toronto) (on file with author).

28 See Terrence F. Ackerman, *An Ethical Framework for the Practice of Paying Research Subjects*, *IRB: ETHICS & HUM. RES.*, July-Aug. 1989, at 1, 2; Ruth W. Grant & Jeremy Sugarman, *Ethics in Human Subjects Research: Do Incentives Matter?*, 29 *J. MED. & PHIL.* 717, 719-23 (2004) (distinguishing between different types of payments).

29 ABADIE, *supra* note 27, at 5, 13, 22-23, 93, 121. For more on subject payment rates, see Neal Dickert & Christine Grady, *Incentives for Research Participants*, in *THE OXFORD TEXTBOOK OF CLINICAL RESEARCH ETHICS*, *supra* note 18, at 386, 394; Ondrusek, *supra* note 27, at 10.

30 Carl Elliott, *Guinea-Pigging*, *NEW YORKER* Jan. 7, 2008, http://www.newyorker.com/reporting/2008/01/07/080107fa_fact_elliott.

31 ABADIE, *supra* note 27, at 75; Trudo Lemmens & Carl Elliott, *Guinea Pigs on the Payroll: The Ethics of Paying Research Subjects*, 7 *ACCOUNTABILITY RES.* 3, 13 (1999); David B. Resnick, *Compensation for Research-Related Injuries*, 27 *J. LEG. MED.* 263, 265 (2006).

drugs may be riskier than others,³² and monitoring for the duration of a study will not catch any problems that manifest later.³³ Unfortunately, there is little empirical data on the overall safety of participation as a healthy research subject.³⁴

As one might expect, like workers, many healthy subjects are motivated to participate in nontherapeutic clinical research primarily by economic factors, even if other reasons also play a secondary role.³⁵ The money offered can be lucrative for relatively short spurts of commitment, leading some healthy subjects to become repeat players,³⁶ although it is unclear precisely how large this subset is. In fact, both the numerator and denominator are relative unknowns because there are several barriers to developing accurate estimates of how many healthy subjects are enrolled in research overall, and how much of the demand for the healthy subjects is met by those who have participated before. There is no national registry of trial subjects, although some sites have developed local registries to address the problem of overlapping enrollment. For example, a group of five sites in Florida found that over the course of 18 weeks in 2009-2010, 2081 individuals attempted to enroll in 27 Phase I studies, 2.4% (50) of whom attempted to enroll in another study within 30 days of receiving a dose in a previous study, and an additional 8.9% (186) attempted to enroll again within 60 days.³⁷ Local registries obviously provide only a limited snapshot, however, and can be circumvented;³⁸ the Florida cohort did not include two large research sites in the area. Thus, U.S. and worldwide numbers remain elusive, but the Florida data alone indicate that the overall number of healthy subjects is not trivial, and that the subset of repeat subjects is also substantial. In fact, there exists a small

32 Roberto Abadie, *The Professional Guinea Pig*, BIOEDGE (Sept. 14, 2010), http://www.bioedge.org/index.php/bioethics/bioethics_article/9204/.

33 ABADIE, *supra* note 27, at 7-8, 82-83.

34 Adil E. Shamoo & David B. Resnick, *Strategies to Minimize Risks and Exploitation in Phase One Trials on Healthy Subjects*, 6 AM. J. BIOETHICS, at W1, W4-W5 (2006).

35 Leanne Stunkel & Christine Grady, *More than the Money: A Review of the Literature Examining Healthy Volunteer Motivations*, 32 CONTEMP. CLINICAL TRIALS 342 (2011); *see also* ABADIE, *supra* note 27, at 5, 41; Ondrusek, *supra* note 27, at 178-79; Olivia Katrandjian, *Growing Number of Mothers Participating in Clinical Trials to Make Ends Meet*, ABC NEWS (Nov. 21, 2011), <http://abcnews.go.com/Health/growing-number-mothers-participating-clinical-trials-make-ends/story?id=14999849>. For a description of the type of people who typically become healthy research subjects, *see* ABADIE, *supra* note 27, at 32; GUINEA PIG ZERO, *supra* note 27, at 212-214; Sheldon Zink, *Maybe We Should Pay Them More*, 1 AM. J. BIOETHICS 88, 88 (2001); Elliott, *supra* note 30.

36 *See* ABADIE, *supra* note 27, at 66; Shamoo & Resnick, *supra* note 34, at W3; Ondrusek, *supra* note 27, at 5.

37 Darran Boyar & Norman M. Goldfarb, *Preventing Overlapping Enrollment in Clinical Studies*, J. CLINICAL RES. BEST PRACTICES 2 (Apr. 2010), http://firstclinical.com/journal/2010/1004_ClinRSVP.pdf; *see also* Giovan Maria Zanini & Claudio Marone, *A New Job: Research Volunteer?*, 135 SWISS MED. WKLY. 315, 315-317 (2005) (reporting similar results from a Swiss register of healthy subjects).

38 David B. Resnick & Greg Koski, *A National Registry for Healthy Volunteers in Phase I Clinical Trials*, 305 JAMA 1236, 1237 (2011).

community of individuals who view themselves as “professional guinea pigs,”³⁹ some of whom have been able to turn participation into their livelihood.⁴⁰

The increasing commercialization of clinical research over the past several decades, particularly at the early phases,⁴¹ combined with the motivations, perspectives, and behaviors of healthy participants in nontherapeutic research has led some to call for regulatory changes that would more accurately reflect this current state of affairs. Trudo Lemmens and Carl Elliott, for example, are troubled by the failure of regulatory and ethical guidelines to recognize that a great deal of research involving healthy subjects is essentially a business transaction. They argue that even though everyone knows better, “researchers have to pretend that [healthy] subjects are motivated by something other than money.”⁴² Research subjects cannot negotiate payment, since payment is not supposed to be the focus of the transaction. Local research ethics boards are expected not to determine what is fair, but what is “undue inducement” per the regulations.⁴³ The result is that healthy subjects face exploitation due to being essentially prohibited “from receiving a fair wage and denie[d] . . . the legal resources available to other high risk workers.”⁴⁴ In particular, healthy subjects “get none of the rights or benefits that come with a good job, such as workers’ compensation, the right to unionize, disability benefits, or health insurance.”⁴⁵

Reasoning that healthy and patient subjects have different interests, vulnerabilities, and motivations for participating, Lemmens and Elliott maintain that it is a mistake not to distinguish between them.⁴⁶ They suggest that a labor

39 See, e.g., ABADIE, *supra* note 27; GUINEA PIG ZERO, *supra* note 27; Helms, *supra* note 5; Ondrusek, *supra* note 27. Note also that the “professional” subject in existence today is a far cry from the ideal form suggested by Maureen Rist and William J. Mohan in *Wanted: Professional Research Subjects; Rewards Commensurate with Risks*, 6 HASTINGS CENTER REP. 28 (1976).

40 Elliott, *supra* note 30. Some subjects rely on research for a substantial proportion of their income, while others use the money as a needed supplement to cover bills or provide extra cash. See ABADIE, *supra* note 27, at 5, 32, 42; Ondrusek, *supra* note 27, at 72-75, 106, 172.

41 Richard Rettig, *The Industrialization of Clinical Research*, 19 HEALTH AFF. 129 (2000); Miriam Shuchman, *Commercializing Clinical Trials – Risks and Benefits of the CRO Boom*, 357 NEW ENG. J. MED. 1365 (2007); Elliott, *supra* note 30.

42 However, federal regulators are beginning to soften on the hard line they have historically drawn regarding financial motives for participating in research. See, e.g., Office of Human Research Protections (OHRP), *Informed Consent: Frequently Asked Questions*, HHS (Mar. 2011), <http://www.hhs.gov/ohrp/policy/consentfaqsmar2011.pdf> [hereinafter OHRP, *Informed Consent*] (noting in response to Question 7 that “compensation may be an acceptable motive for agreeing to participate in research”).

43 Trudo Lemmens & Carl Elliott, *Justice for the Professional Guinea Pig*, 1 AM. J. BIOETHICS 51, 52 (2001).

44 *Id.* at 53.

45 Carl Elliott & Roberto Abadie, *Exploiting a Research Underclass in Phase I Clinical Trials*, 358 NEW ENG. J. MED. 2316, 2317 (2008).

46 Savulescu also takes issue with the payment structure of clinical research, but without distinguishing between healthy and patient subjects. Instead, he distinguishes between commercial and noncommercial research, arguing that subjects should be offered “fair compensation and

model is more appropriate for paid healthy subjects research, which takes place in a commercial sphere where the research institution and subject enter the relationship on financial grounds in the absence of shared interests.⁴⁷ Indeed, they suggest that “the regulatory regime and protective measures offered by many labor and occupational health and safety laws may provide a very useful model for the protection of subjects” in a number of ways:

For example, labor-type legislation could empower occupational health and safety agencies to conduct inspections, and it could lead to the establishment of occupational health and safety committees in which subjects are represented Collective negotiations and unionization would be a way of empowering subjects and of giving them a stronger voice in arguing for good working conditions. As with workers’ compensation, appropriate compensation schemes could be enforced to offer some form of financial security in case subjects are harmed in research. Standards of remuneration could be negotiated, based on the level of discomfort, the number and types of procedures, the duration of the studies, particular circumstances such as the obligation to remain in the research institution for a lengthy period of time, and so on. Regulation could also require . . . that research subjects have disability insurance and other insurance coverage to protect them in cases of “occupational injury.”⁴⁸

Lemmens and Elliott claim that action is needed on these fronts, since research participation is not considered “employment” at present, leaving subjects ineligible for the resources and protection provided by existing worklaws.⁴⁹ However, they are careful to clarify that their linkage between healthy subjects research and labor contracts is not intended to be literal, emphasizing that “research has particular characteristics which warrant treating it as a category *sui generis*.”⁵⁰ They also raise some additional concerns suggesting that they do not fully buy into the analogy, not all of which are convincing. For example, they worry about the risk of further commercializing healthy subjects research, exacerbating the risk of exploitation if restrictions on payment are

salary” when they participate in a project intended to bring in profit. Julian Savulescu, *On the Commercial Exploitation of Participants of Research*, 23 J. MED. ETHICS 392 (1997).

47 Lemmens & Elliott, *supra* note 31, at 4-5, 14-16. Many authors have argued for differential treatment of healthy subjects and patient subjects. See Lemmens & Elliott, *supra* note 43, at 52; Ari VanderWalde, *Undue Inducement: The Only Objection to Payment*, 5 AM. J. BIOETHICS 25, 26-27 (2005); Ondrusek, *supra* note 27, at 197.

48 Lemmens & Elliott, *supra* note 31, at 16-17.

49 Lemmens & Elliott, *supra* note 43, at 52.

50 Lemmens & Elliott, *supra* note 31, at 16.

removed without adding other protections (particularly for safety), and potentially hindering important research that cannot afford to pay subjects a competitive price.⁵¹ Several of these issues are taken up in Part IV.A.

Other bioethicists have taken a similar stance with regard to advocating for certain employment-type protections of healthy Phase I research subjects, but have gone even further. For example, Adil Shamoo and David Resnick also argue against payment limits for healthy subjects,⁵² and suggest that these subjects, “like all workers,” should receive compensation for injury and be permitted to use collective bargaining to negotiate for higher pay and better working conditions. Unlike Lemmens and Elliott, however, Shamoo and Resnick seem to have no reservations regarding the wholesale adoption of a labor approach to healthy subjects research. The same is true of Roberto Abadie, who rather than drawing a mere analogy to work, suggests that in order to avoid unacceptable exploitation we “need to recognize that [healthy] volunteers’ participation *is* labor, even if it is what they call a ‘weird type of work,’ and provide better working conditions and proper compensation.”⁵³ “Paid subjects,” he maintains, “should be given the same labor protections guaranteed to other workers in risky occupations.”⁵⁴

2. Equating Subject Payment with a Wage

Exploration of work-type protections for HSR has also been prompted by a different starting point, a widely discussed proposal by Neal Dickert and Christine Grady to treat payment to research subjects as a wage. Recognizing that paid participation raises a variety of ethical concerns, but also that payment is common, probably necessary in some cases, and currently without adequate regulatory guidance, Dickert and Grady set forth three models on which payment to healthy and patient-subjects could conceivably be based.⁵⁵ After rejecting both a market model, in which supply and demand for subjects dictate how much they are paid, and a reimbursement model, in which payment is limited to either reimbursement of a subject’s actual expenditures or his time away from work, Dickert and Grady ultimately endorse a wage-payment model on the basis that research participation is analogous to other essential but unskilled labor, and should be paid on a commensurate scale. While recognizing some drawbacks, they favor this as the best option because it would reduce the possibility that subjects will be unduly induced to participate against their better judgment, as

51 Lemmens & Elliott, *supra* note 43, at 52-53; see also Neal Dickert & Christine Grady, *What’s the Price of a Research Subject? Approaches to Payment for Research Participation*, 341 NEW ENG. J. MED. 198, 199-201 (1999) (rejecting the market model).

52 Shamoo & Resnick, *supra* note 34, at W9-W11 (2006); see also VanderWalde, *supra* note 47, at 27.

53 ABADIE, *supra* note 27, at 165-66 (emphasis added).

54 *Id.*

55 Dickert & Grady, *supra* note 51, at 198.

well as reduce the financial sacrifice required of subjects, while allowing them to be paid for work that is valuable to society. In addition, this approach would set a lower limit on the amount paid with all the associated benefits of a minimum wage, and the standardization of pay across studies would result in less competition for subjects between them.⁵⁶

Despite this conclusion, however, Dickert and Grady did not suggest that a labor approach to research is appropriate across the board. James Anderson and Charles Weijer, on the other hand, have considered the moral implications of treating subjects as wage-earners or entrepreneurs on a broader scale.⁵⁷ They note that *if* participation is a job for which one may be paid, morally indistinguishable from other lines of legitimate employment, then similar cases must be treated similarly,⁵⁸ and the ethics of research participation must be viewed through the same moral lens used to determine just working conditions in other contexts.⁵⁹

Anderson and Weijer suggest that management's obligations to workers vary depending on whether the workers are part-time or full-time, temporary or "career," and that these distinctions are mirrored among human subjects, who may be characterized as either "occasional" or "professional."⁶⁰ They go on to explore the potential features of just working conditions, including those minimum protections set by law, as well as additional normative requirements such as the right to meaningful work and the fair distribution of "hard work."⁶¹ Anderson and Weijer recognize that each of these rights does not necessarily apply to all workers, but rather depends on their level of dependency, commitment, and investment.

Thus, like both temporary and career workers do, both occasional and professional subjects have a right to at least a minimum wage.⁶² Both are protected by standard work week legislation and have a right to overtime pay for hours outside of that standard. Both have a right to a safe workplace, and no fault compensation for work related injury. And both have the right to organize into unions. Similarly, professional subjects, like career workers, [also] have a right to a pension, to benefits such as

56 *Id.* at 199-201.

57 James A. Anderson & Charles Weijer, *The Research Subject as Entrepreneur*, 1 AM. J. BIOETHICS 67 (2001) [hereinafter Anderson & Weijer, *Entrepreneur*]; James A. Anderson & Charles Weijer, *The Research Subject as Wage Earner*, 23 THEORETICAL MED. & BIOETHICS 359 (2002) [hereinafter Anderson & Weijer, *Wage Earner*].

58 Anderson & Weijer, *Wage Earner*, *supra* note 57, at 360.

59 Charles Weijer, *Meaningful Work as Due Inducement*, 26 THEORETICAL MED. & BIOETHICS 431-32 (2005).

60 Anderson & Weijer, *Wage Earner*, *supra* note 57, at 361-62.

61 *Id.* at 364-70.

62 *Id.* at 375. Anderson and Weijer appropriately take issue with Dickert and Grady's suggestion that subject payment should be kept low.

medical and dental insurance, and finally, they have the right to meaningful work.⁶³

Of course, at present, subjects are not treated this way, but rather “many of the deplorable working conditions characteristic of the industrial revolution are being duplicated in the realm of biomedical research.”⁶⁴

This disparity is problematic and unjust, Anderson and Weijer maintain, *if subjects are workers too*.⁶⁵ Thus, we must either adjust present practice or reject the idea of paid subjects as unskilled wage-earners.⁶⁶ They do not explicitly state which approach they favor, although they do suggest some discomfort with treating research subjects like other workers and indicate that the morally satisfactory solution to the issue of subjects-for-hire requires something more than improved working conditions or unionization.⁶⁷ Others, however, maintain that “paid research subjects must have the same rights, benefits, and payment schemes available to all workers of similar type.”⁶⁸

3. Analysis

Although the bioethics literature considering the analogy between human subjects and more traditional workers has certainly moved the ball forward, the arguments remain incomplete. This is understandable given that the existing literature begins from a different starting point, using the analogy to identify solutions to imperfect subject protections, whereas this Article is more interested in the analogy for its own sake to identify inconsistencies in need of resolution. So what work remains?

First, while there has been some attempt to differentiate between different types of subjects and recognition of the fact that different types of workers generally get different rights and benefits, the distinctions that have been drawn by references to high-risk work, “good” jobs, and temporary versus career work do not necessarily map onto the lines drawn by the law, which predominantly distinguishes between independent contractors and employees. Many have criticized these legal lines and how they are drawn, as described below, but the fact remains that there are entire categories of workers who stand without the protections under discussion for subjects. The case has not yet been adequately made as to why subjects are distinguishable from independent contractors, who do not receive these protections, and analogous to employees, who do.

63 *Id.* at 371.

64 *Id.* at 372.

65 Anderson & Weijer, *Entrepreneur*, *supra* note 57, at 68.

66 Anderson & Weijer, *Wage Earner*, *supra* note 57, at 374.

67 Anderson & Weijer, *Entrepreneur*, *supra* note 57, at 68; Weijer, *supra* note 59, at 432.

68 VanderWalde & Kurzban, *supra* note 26, at 555.

Another aspect of the debate that is far from resolved is whether human subjects actually engage in work. A number of potentially relevant criteria have been raised to date, including payment, and more specifically payment as a wage or something else;⁶⁹ subject motivation and source of vulnerability;⁷⁰ acting versus being acted upon;⁷¹ commodification of the body itself compared to sale of a service;⁷² risk and uncertainty as key features of the activity or mere byproducts;⁷³ and the perceptions of relevant players.⁷⁴ Resolution of this question in favor of participation as work would certainly strengthen the argument regarding subject protection, at least as a matter of consistency. However, given the philosophical complexity of accurately defining work, and because the argument regarding subject protection can also be made on the basis of analogy and shared goals and challenges, rather than on a definitional basis, the ontological question will not be further considered here.⁷⁵

But note that some have even rejected the very *analogy* that would compare research participation to work, or have at least questioned the analogy's application. For example, when commissioned for comment on the issue by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, philosopher Marx Wartofsky argued that what is being sold by the research subject is different than what is being sold by the worker in a hazardous occupation; "[t]he coal miner is paid to mine coal," he claimed, but the subject is "paid to put himself at risk in order to see what the effects will be on his body or his health."⁷⁶ This distinction fails, however, because in both cases individuals are intentionally exposed to the risk of harm in order to achieve some other goal. More simply, the subject could just as easily be seen as paid to provide data – the primary purpose of HSR – as to accept risk, rendering risk a byproduct rather than a defining feature of research.

On the other hand, for certain types of research, such as challenge studies involving intentional exposure to infection and drug safety studies in healthy subjects, and for many types of research procedures done exclusively to gather research data, the primary goal of generating generalizable knowledge can only be achieved through the means of intentionally inflicting harm, rather than merely foreseeing it as a side effect. This is different from nearly every other

69 See, e.g., Lemmens & Elliott, *supra* note 31; Anderson & Weijer, *Wage Earner*, *supra* note 57; Grant & Sugarman, *supra* note 28.

70 See, e.g., Lemmens & Elliott, *supra* note 31, at 14-15; Shamoo & Resnick, *supra* note 34.

71 See, e.g., Lemmens & Elliott, *supra* note 31; Wartofsky, *supra* note 7.

72 See, e.g., Dickert & Grady, *supra* note 51; Lemmens & Elliott, *supra* note 31, at 16; Wartofsky, *supra* note 7.

73 See, e.g., Lemmens & Elliott, *supra* note 31, at 16; Wartofsky, *supra* note 7.

74 See, e.g., Lemmens & Elliott, *supra* note 31.

75 I do hope to address it in future scholarship, however. For now, suffice it to say that if prostitution and modeling can be considered work, participation as a human research subject can be as well. See GREGOR GALL, *SEX WORKER UNION ORGANISING: AN INTERNATIONAL STUDY*, 23-34 (2006).

76 Wartofsky, *supra* note 7, at 3-11 to 3-12.

type of risky work in which harm may be foreseen, but is neither intended nor useful toward any goal, and in fact, hopefully avoided.⁷⁷ This distinction may be important with regard to justifying the special *extra* protections that are extended to research subjects, such as IRB review and informed consent. However, it certainly does not suggest that subjects should have fewer protections than workers or that it is inherently problematic or offensive to analogize between the two.

Nonetheless, Paul McNeill also maintains that risky work is a poor analogy to HSR, arguing that dangerous work “must be justified by some greater need” and “every effort should be taken to minimize the risks inherent in the work.”⁷⁸ Although this is precisely what is required of HSR, and in fact not required in practice or by law for many types of work (such as crab fishing or diamond mining), McNeill questions whether research actually needs to be done at all, since progress is an “optional goal” and the “justification for research is always in terms of progress.”⁷⁹ However, if dangerous work such as fire fighting is necessary, which McNeill concedes, why is dangerous work such as research participation – which may also save lives and meet basic human needs – any less so? There seems to be no reason to distinguish between different types of potentially preventable deaths when people have voluntarily put themselves at risk in the service of a greater good. Terrence Ackerman also tries to distinguish research from risky work, arguing that the “roles performed by police officers, fire fighters, and soldiers are absolutely essential to the welfare of society. . . . By contrast, conduct of clinical research involving serious dangers to the welfare of subjects (which they would otherwise avoid) is generally not necessary to achieve the goals of medical research.”⁸⁰ If serious dangers to subject welfare in any particular study were not actually necessary to the research, however, IRBs would by regulation be required not permit them. Thus, Ackerman’s distinction fails as well.

There may in fact be some real differences between at least certain types of research participation and traditional work, but none of these differences suggest that worklaw protections would be inappropriate in the research context, and the analogy is strong enough on its face to sustain serious comparative analysis. The bioethicists who have put the analogy forth to date, however, have not adequately considered the goals and rationales of worklaw and whether they overlap with those of human subjects protection. This is primarily because the bioethicists have identified some problematic aspects of the existing human subjects protection regime and seek lessons from other areas as to how things could be improved. In that context, all one needs to know is whether worklaw approaches

77 Thank you to Collin O’Neil for pointing out this distinction.

78 McNeill, *supra* note 7, at 392.

79 *Id.*

80 Ackerman, *supra* note 28, at 4.

would in fact offer such improvement. But again, the problem motivating this Article is not some concern that the HSR regulations are inadequate, but rather concern that the HSR regulations are *inconsistent*. And for that type of analysis, a deeper understanding of why certain workers are protected in certain ways is essential to demonstrating that such inconsistency ought to be resolved.

B. Legal Background

Before getting to the question of “why,” however, we must address some foundational questions regarding the type of worker most appropriate for comparison to human subjects.

1. Employees and Employers

Whether at the federal, state, or local level, the laws that regulate working relationships generally cover employees only, as opposed to independent contractors or volunteers.⁸¹ Thus, non-employees are usually ineligible for wage and hour protections, unemployment benefits, workers’ compensation, family and medical leave, protected rights to form or join a union and collectively bargain, anti-discrimination and anti-retaliation protections, and workplace safety protections, for example. But who counts as an employee?

Worklaw statutes themselves often contain a number of specific inclusions and exclusions that help clarify whether a given worker can claim coverage. With regard to the definition of employee itself, however, statutory language is notoriously tautological. For example, the Occupational Safety and Health Act covers “employees,” but defines the term as “an employee of an employer who is employed in a business of his employer which affects commerce,” and an “employer” as a “person engaged in a business affecting commerce who has employees”⁸² Similar circularity is found in the National Labor Relations Act,⁸³ and a number of other federal and state laws.⁸⁴ As a result, courts and administrative agencies have had to develop a number of tests for distinguishing employees from other types of workers.

One way to defeat employee status is to find that the work in question is beyond the economic or market sphere, such as that done in the home or by prisoners.⁸⁵ Another is to determine that the worker is a volunteer because he

81 Richard R. Carlson, *Why the Law Still Can't Tell an Employee When It Sees One and How It Ought to Stop Trying*, 22 BERKELEY J. LAB. & EMP. L. 295, 366 (2001); Noah D. Zatz, *Working Beyond the Reach or Grasp of Employment Law*, in THE GLOVES-OFF ECONOMY: WORKPLACE STANDARDS AT THE BOTTOM OF AMERICA'S LABOR MARKET 31, 34 (Annette Bernhardt et al. eds., 2008).

82 29 U.S.C. § 652(5)-(6) (2006).

83 29 U.S.C. § 152(2)-(3) (2006).

84 Carlson, *supra* note 81, at 296, n.5.

85 Noah D. Zatz, *Working at the Boundaries of Markets: Prison Labor and the Economic Dimension of Employment Relationships*, 61 VAND. L. REV. 857, 861-62 (2008).

receives – and expects to receive – no financial compensation or benefits in return for his work.⁸⁶ The receipt of certain benefits may push a worker out of volunteer territory,⁸⁷ but mere reimbursement of expenses incurred incidental to the work is likely insufficient to do so.

Employee status also may be defeated by demonstrating that a worker is an independent contractor, but unfortunately, there is substantial uncertainty and disagreement regarding exactly how and when to make this distinction. The precise rules and definitions provided in statutes, regulations, and judicial opinions are usually both complex and vague.⁸⁸ Courts have relied on a number of different approaches, including the common law “direct and control test” (which asks whether the hiring party retains the right to direct and control “not only what shall be done, but how it shall be done”),⁸⁹ the “economic realities test” (which examines the whole activity of the relationship to determine the extent of the hiring party’s domination over the worker and the worker’s economic dependence);⁹⁰ and the “statutory purpose test” (which defines employees “in light of the mischief to be corrected and the end to be attained” by the law in question).⁹¹ A great deal has been written about the similarities and differences between each of these approaches,⁹² and certain tests have been adopted over others in the context of interpreting particular statutes. The important thing, however, is that in nearly every jurisdiction and for nearly every worklaw, the relevant test of employee status involves an intricate weighing of a laundry list of case-specific factors, which often point – or can be manipulated – in different directions.⁹³ As a result, these tests, which apply precisely as a result of statutory ambiguity, are often ambiguous themselves, leaving substantial room for unpredictability, confusion, and differences of opinion about the status of

86 Mitchell H. Rubinstein, *Our Nation’s Forgotten Workers: The Unprotected Volunteers*, 9 U. PA. J. LAB. & EMP. L. 147, 154 (2006).

87 *Id.* at 157; Zatz, *supra* note 85, at 920 n.293.

88 Zatz, *supra* note 81, at 35.

89 *Singer Mfg. Co. v. Rahn*, 132 U.S. 518, 523 (1889); see also Carlson, *supra* note 81, at 304-306, 309-11; Katherine V.W. Stone, *Legal Protections for Atypical Employees: Employment Law for Workers Without Workplaces and Employees Without Employers*, 27 BERKELEY J. EMP. & LAB. L. 251, 280 (2006); Zatz, *supra* note 81, at 35.

90 Carlson, *supra* note 81, at 311-13, 314, 317, 327, 342; Stone, *supra* note 89, at 280; Zatz, *supra* note 81, at 35.

91 *N.L.R.B. v. Hearst*, 322 U.S. 111, 124 (1944) (quoting *South Chicago Coal & Dock Co. v. Bassett*, 309 U.S. 251, 259 (1940)). For a discussion of the current status of the statutory purpose test, compare Carlson, *supra* note 81, at 333, 338, 343, 353 and Zatz, *supra* note 81, at 35 (suggesting that the test is essentially dead as a matter of determining federal statutory coverage), with Kenneth G. Dau-Schmidt & Michael D. Ray, *The Definition of “Employee” in American Labor and Employment Law*, JAPAN INST. LAB. POL’Y & TRAINING 124, http://www.jil.go.jp/english/events_and_information/documents/clls04_dauschmidt2.pdf (last visited Nov. 21, 2013) (indicating that in most areas of worklaw, status as an employee is defined according to the purposes of the statute).

92 See, e.g., Carlson, *supra* note 81, at 309-44.

93 *Id.* at 343.

workers who seem to fall somewhere between employee and independent contractor (or volunteer).⁹⁴

It is also important to recognize that there is a second component to determining a worker's employment status: identifying which party, if any, will ultimately be responsible for providing the benefits or protections afforded to employees under various laws. On the upside, the law recognizes joint employment relationships in which a single employee can simultaneously have multiple employers for the purposes of particular work.⁹⁵ However, there is also a growing phenomenon of "employerless employees"⁹⁶ and "disintegrating employers."⁹⁷ When there are multiple parties involved, the features of an employment relationship that usually hang together may be split up.⁹⁸ As a result, each of the putative employers can plausibly deny employer status on the grounds that none individually satisfies the relevant legal criteria. Thus, a worker who does not really control his own work without dependence on any other party may be left without any employer from whom he could demand benefits.

2. *The Employment Status of Human Research Subjects*

Some of the shortcomings of worklaw categories and tests are apparent even from consideration of the limited instances in which the status of human subjects has been given direct legal attention. For example, in one of the more interesting opinions issued to date, an unemployment board of appeals relied on the intervention of a contract research organization (CRO) between the company seeking data and subjects participating in drug absorption trials to find that the subjects seeking benefits were independent sub-contractors, rather than employees.⁹⁹ Importantly, CROs, which can take on any or all of a sponsor's regulatory obligations, and perform services ranging from recruitment to actually running trials in their own dedicated facilities, are increasingly being utilized to carry out clinical studies quickly and efficiently,¹⁰⁰ rendering subjects' employment status even more precarious.

The state unemployment law in question defined workers as independent contractors ineligible for coverage when they satisfied a three-pronged test,¹⁰¹

94 Stephen D. Befort, *Revisiting the Black Hole of Workplace Regulation: A Historical and Comparative Perspective of Contingent Work*, 24 BERKELEY J. EMP. & LAB. L. 153, 167 (2003); Carlson, *supra* note 81, at 301, 335-53.

95 Zatz, *supra* note 81, at 41.

96 See generally Stone, *supra* note 89.

97 Zatz, *supra* note 81, at 37.

98 Richard R. Carlson, *Variations on a Theme of Employment: Labor Law Regulation of Alternative Worker Relations*, 37 S. TEX. L. REV. 661, 688-91 (1996).

99 Board of Appeals Decision, [Md.] Unempl. Ins. Rep. (CCH) ¶ 8469 (Nov. 14, 1994), 1994 WL 1686552.

100 Phillip Mirowski & Robert Van Horn, *The Contract Research Organization and the Commercialization of Scientific Research*, 35 SOC. STUD. SCI. 503, 509-13 (2005).

101 MD. CODE. ANN., LAB. & EMPL. § 8-205 (West 1991).

which the Board determined to be met here. First, the Board held that none of the control exerted by the CRO, Pharmakinetics, over the subjects was adequate to create an employer-employee relationship. The Board explained that when a general contractor hires a sub-contractor, the sub-contractor is not entirely unrestricted in his work, but rather must follow the set of plans established by the customer. Because these restrictions are imposed by the customer, however, they do not constitute an employer's control by the general contractor, even when he takes steps to ensure they are met.¹⁰² In the research context, then, the Board construed the drug company as the customer, the research protocol as the customer's plans, Pharmakinetics as the general contractor, and the subject as the sub-contractor.¹⁰³ It also suggested that because after the subject ingests the drug, "it is *his/her body which is actually performing the work*," the general contractor cannot "logically" have control over the performance of the subject's work and can rather only monitor for proper administration.¹⁰⁴

With regard to the second prong, customary engagement in the independent performance of the work involved, the Board maintained that the subjects were "engaged in the business of providing drug testing services," and emphasized that they may be involved in multiple studies for multiple companies.¹⁰⁵ Third and finally, the Board had to decide whether the subjects' work was outside the usual course of Pharmakinetics' business, which it determined to be analyzing data on drug absorption and providing results to companies; "[t]esting drugs is merely the part of their business in which they derive their data."¹⁰⁶ On the other hand, it found that the work of the subjects is to test the drugs – to ingest them and provide bodily fluids for analysis.¹⁰⁷ The subjects' work was clearly integral to Pharmakinetics' business, but the Board held it was nonetheless outside its scope, arguing that just as Pharmakinetics needs subjects, "the presence of a subcontractor plumber *is needed and necessary in the course of business of his/her general contractor*."¹⁰⁸

Although it emphasized the general contractor/sub-contractor analogy, it is noteworthy that at least the second and third prongs of the Board's (questionable) analysis would similarly apply had a drug company engaged the subjects for a test they were running themselves without the aid of a CRO. Nonetheless, the Board's analogy between a protocol and a customer's plans seems flawed. Plans may specify certain standards and aspects of process, but they would not specify precise details of how the work must be done, focusing attention instead on the end result. Clinical trial protocols, on the other hand, are of course concerned

102 Board of Appeals Decision, *supra* note 99, at 3.

103 *Id.* Note, however, that this case might have reasonably been decided as an example of joint employment.

104 *Id.* at 4.

105 *Id.*

106 *Id.*

107 *Id.* at 4-5.

108 *Id.* at 5.

with the end result – the data – but are also enormously detailed with regard to schedules, procedures, and other restrictions, all features strongly indicative of an employer’s type of control. Ultimately, the Board tried to draw a distinction between managing the integrity of the work (customer/contractor) and controlling its performance (employer/employee).¹⁰⁹ However, the two are really one and the same in the clinical trials context.

In contrast to the Pharmacokinetics decision, consider *Edward Lowe Industries, Inc. v. Missouri Division of Employment Security*,¹¹⁰ which reached a different conclusion on a quasi-subject’s employment status for purposes of unemployment benefits. Lowe engaged the services of paid consumer panelists, including the claimant Zeta Simms, to conduct studies evaluating the odor of different cat litter formulations.¹¹¹ The court considered the factors set forth by the IRS to flesh out the common law control test, and determined that on the totality of the circumstances, Simms was in fact an employee.¹¹² However, the state unemployment law explicitly stated that “the term ‘employment’ shall not include . . . [s]ervices performed as a volunteer research subject who is paid on a per study basis for scientific, medical or drug related testing”¹¹³ This particular statutory carve out of certain types of research subjects as non-employees is an interesting indication of policy choice regarding which individuals ought to receive unemployment protection, but in this case, the court concluded that Simms’ services did not fall within the exclusion on the grounds that the litter tests were clearly not medical or drug related, nor were they particularly scientific.¹¹⁴

Notably, the I.R.S. has applied its factors directly to human subjects as well, reaching the same conclusion on employment status as the court in *Lowe*.¹¹⁵ Other legal decision makers have also determined that subjects can be employees. For example, when Qualia Clinical Services declared bankruptcy, research subjects who had not been paid as promised filed claims under a priority wage provision of the Bankruptcy Code. In response to Qualia’s challenge, the court held that entitlement to priority wage claims could extend even to individuals who are not employees depending on whether “something more than a ‘mere contractual relationship’ existed between the parties, as well as on the

109 *Id.* at 3.

110 865 S.W.2d 855, 856 (Mo. Ct. App. 1993).

111 *Id.* at 858-59.

112 *Id.* at 860-63.

113 *Id.* at 863 (quoting MO. ANN. STAT. § 288.034(12)(17) (West 1993) (current version at § 288.034(12)(18) (West 2013)). Other states have nearly identical provisions. *See, e.g.*, MD. CODE ANN., LAB. & EMPL. § 8-222 (West 1995); NEB. REV. STAT. § 48-604(6)(s) (2011).

114 865 S.W.2d at 863.

115 I.R.S. Priv. Ltr. Rul. 92-34-024 (Aug. 21, 1992) (determining that a paid subject participating in nutritional studies while living at a research center was an employee for purposes of tax withholding requirements).

amount of control exercised over the claimants by the debtor.”¹¹⁶ Because subjects had signed consent forms that were lengthy and detailed, the study protocols required rigid control, studies were conducted at Qualia’s site, and the subjects were overseen by Qualia staff, the court determined that the “extent of Qualia’s control over the claimants’ performance in the studies effectively rendered them ‘employees.’”¹¹⁷ Qualia’s role as a CRO seems to have played no part in the analysis.

Research subjects have also claimed employee status in the context of sexual harassment and minimum wage/overtime claims, although it is unclear whether they would have been successful because their cases settled out of court.¹¹⁸ Further, the Comptroller General has weighed in on this question, at least tangentially. In 1966, the Army sought to use civilians as human subjects in its experiments, secured under contracts on a fee basis.¹¹⁹ However, a government rule allowed such contracts only if the service would be performed without detailed control over the method by which the desired result would be accomplished.¹²⁰ The Army explained that this condition was satisfied because it was “not interested so much in what the subject produces through his efforts, as in measuring and examining the subject’s reactions to a set of conditions induced by the government. The government exercises no control over the subject’s reactions in the sense of directing the subject how to react, and in fact the experimentation is valid only if his reactions are purely independent and objective.”¹²¹ The Comptroller was convinced that the situation did not “clearly fall within the rules for establishing an employee-employer relationship,”¹²² and saw no objection to treating the subjects as independent contractors without entitlement to leave, retirement, salary, tenure, etc.¹²³

116 *In re Qualia Clinical Servs., Inc.*, No. BK09-80629-TJM, 2009 WL 2513820, at 1 (Bankr. D. Neb. Aug. 11, 2009).

117 *Id.* at 2.

118 Email from Scott Magaw, attorney for the plaintiffs (Feb. 12, 2012)(on file with author). The cases in question were *Krah v. Univ. of Pitt.*, No. 2:97-cv-00834-DEZ (W.D. Pa. 1997), and *Cortazzo v. Univ. of Pitt.*, No. 2:97-cv-00832-DEZ (W.D. Pa. 1997). See also Hank Grezlak, *Research Participants Sue for Sexual Harassment*, PENN. L. WKLY., May 5, 1997; *Doc Watch: Lady Lab Rats Sue Researcher*, 4 GUINEA PIG ZERO 27 (article on file with author).

119 To the Sec’y of the Army, 45 Comp. Gen. 649, 649 (1966).

120 *Id.* at 650.

121 *Id.* at 649.

122 *Id.* at 650.

123 *Id.* at 650-51; see also SEC’Y OF THE AIR FORCE, PROTECTION OF HUMAN SUBJECTS IN BIOMEDICAL AND BEHAVIORAL RESEARCH, AIR FORCE INSTRUCTION NO. 40-402 (May 5, 2005) (noting that private individuals participating as research subjects may enter into an independent contractor relationship with the Air Force).

Beyond this handful of scattered examples, however, which present arguments and outcomes on both sides of the line, there is virtually no discussion of subjects' employment status in the legal literature.¹²⁴

3. Analysis

The cases considering the employment status of human research subjects present a very small *n*, which is further limited by the fact that none of the cases seems to have involved unpaid research participation, therapeutic research, or patient-subjects, nor does any address issues likely to be of importance to subjects such as workers' compensation. Nonetheless, it is possible to draw a few broad conclusions.

First, despite the unique (or at least unusual) nature of the workers in question, all of the cases involved rigid application of the relevant law, without real consideration of statutory purpose, whether the subjects actually needed the protection they sought, fairness, or possible consequences. This sort of strict, formalist approach is the norm in worklaw,¹²⁵ but stands in stark contrast to the policy focus of the bioethicists described above.

Second, these examples indicate that at least some types of human subjects are likely to have little difficulty with the preliminary hurdles of establishing their employee status, namely the performance of work and existence of an economic relationship. Indeed, with the exception of *Pharmakinetics*' point regarding the subject's body doing the work, none of the decision makers in these cases stumbled over these questions, instead simply assuming the subjects were engaged in market work and focusing exclusively on distinctions between employees and independent contractors. This is notable considering that the prevailing approach in bioethics and research compliance is to treat human subjects as though they are not engaged in work at all.

Finally, these cases suggest that legal avenues of subject protection beyond the current HSR regulations may have been under-pursued to date. That is to say, contrary to Lemmens and Elliott's point that human subjects are not considered employees, at least some legal decision makers have been willing to recognize them as such. Thus, it seems plausible that certain types of subjects seeking additional protections could conceivably achieve them even under existing law, depending on the language of the statute in question, any specific exemptions or

124 For two examples of pieces that very briefly consider whether the exclusion of women from research participation could constitute illegal employment discrimination, see R. Alta Charo, *Protecting Us to Death: Women, Pregnancy, and Clinical Research Trials*, 38 ST. LOUIS U. L.J. 135, 156 (1993); Vanessa Merton, *The Exclusion of Pregnant, Pregnable, and Once-Pregnant People (a.k.a. Women) from Biomedical Research*, 3 TEX. J. WOMEN & L. 307, 358 n.193, 369 (1994).

125 See Guy Davidov, *The Reports of My Death Are Greatly Exaggerated: "Employee" as a Viable (Though Overly-Used) Legal Concept* 13-15 (May 2005) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=783484.

inclusions, and interpretation of each element in the relevant test of employment, as well as the particular details of how the research is carried out.

There is room for disagreement, but the strongest case for employee status seems to exist for paid subjects in nontherapeutic research who are motivated by money, ideally in the absence of a potentially complicating CRO. Those engaged in therapeutic research may have more difficulty establishing the market nature of their work, although this would not necessarily be insurmountable. And the weakest case exists for those who expect to receive nothing in return for their participation in research, since they are clearly volunteers under the law. For reasons that will be explored in the next Part, however, this sort of purely doctrinal approach leaves much to be desired.

III. CLARIFYING THE ANALOGY AND ITS SCOPE

Legal analysis indicates that the “human research worker” *analogy* might be closer to *reality* for some subjects, namely those likely to fit the bill as employees. However, this does not necessarily mean that direct application of existing worklaw is the best approach. Nor does it mean that the analogy fails for other types of subjects, or that they should not also be granted additional protections based on those available to workers. Rather than trying to map different types of subjects directly onto the legal standards for employees, independent contractors, and volunteers, it is preferable to think carefully about precisely which protections to incorporate into the HSR regulations, and how they should be refined for that context

A. Why Not Apply Worklaw Directly to Subjects?

A number of features of worklaw suggest that it is best used only as a model for HSR, rather than a perfect fit. Most critically, many have suggested that existing laws and interpretations are ill-suited to the modern work environment – and if they are in need of modification on their home turf, there is compelling reason to worry about their application to HSR.

Part of the problem is anachronistic in the sense that the employee/independent contractor distinction that is now used to allocate responsibility for a variety of worker benefits and protections was originally developed from the master/servant doctrine of agency law for an entirely different purpose: allocating responsibility for a worker’s tortious conduct. As often happens when applying rules from one situation to another, marginal cases can become difficult to handle. Moreover, as Katherine Stone explains,

[t]he legal rules governing collective bargaining and individual employment rights, as well as the provision of social welfare benefits all assumed the existence of a stable, on-going relationship between an individual and a firm. Now, as firms are

breaking apart, downsizing, rearranging their functions, and dispersing their facilities, they no longer offer the kind of stable long-term relationship upon which our legal rules depend.¹²⁶

Importantly, these concerns are rooted in features of changing working relationships that are inherent in HSR, such as temporary engagement and work for several different parties. In fact, one feature of the modern workplace that has been cited as a challenge to existing legal frameworks – outsourcing¹²⁷ – has an obvious corollary in the conduct of medical research: CROs. And just as traditional workers can problematically be left unprotected as employerless employees, when the components of research are disaggregated across a number of parties in a variety of combinations, the question of who is controlling the subject becomes substantially more complicated.

Those critical of the current state of worklaw do not necessarily take issue with the fact that not all workers are covered.¹²⁸ But the problem for many commentators is that the lines drawn can result in workers who seem to need protection being left without it.¹²⁹ As a result, some have advocated for completely discarding the labeling distinctions between different types of workers,¹³⁰ and instead simply examining whether a particular worker needs protection.¹³¹ Others have suggested experimenting with legal approaches that do not rely so heavily on employer-employee relationships as the source of worker protection;¹³² recognizing new intermediate labels, such as the “dependent contractor” and “uncontrolled employee”;¹³³ simply changing the basic definitions of employment to better encompass those workers sharing the characteristics that justify protection;¹³⁴ and paying closer attention to whether entitlements that have been tied to the existence of an employment relationship really ought to be.¹³⁵

The point here is not to fully engage with the criticisms of current legal approaches, but rather to note their existence and plausibility, which suggest serious concerns about the protections available to workers in general. If we

126 Stone, *supra* note 89, at 254.

127 *Id.* at 253-56; Zatz, *supra* note 85, at 860-61; Zatz, *supra* note 81, at 43.

128 Davidov, *supra* note 125, at 2.

129 Befort, *supra* note 94, at 168; Carlson, *supra* note 81, at 367; Stone, *supra* note 89, at 256-270, 279; Dau-Schmidt & Ray, *supra* note 91, at 117, 120.

130 *But see* Davidov, *supra* note 125 (arguing against this approach).

131 Carlson, *supra* note 81, at 301; Marc Linder, *Dependent and Independent Contractors in Recent U.S. Labor Law: An Ambiguous Dichotomy Rooted in Simulated Statutory Purposelessness*, 21 COMP. LAB. L. & POL’Y J. 187 (1999); Stone, *supra* note 89, at 284.

132 Zatz, *supra* note 81, at 32, 57.

133 Befort, *supra* note 94, at 172-74; Guy Davidov, *Who is a Worker?*, 34 INDUS. L.J. 57, 61-62 (2005); *see also* Linder, *supra* note 131; Stone, *supra* note 89, at 279.

134 Zatz, *supra* note 81, at 51; Davidov, *supra* note 125, at 9, 11.

135 ALAIN SUPLOT, *BEYOND EMPLOYMENT: CHANGES IN WORK AND THE FUTURE OF LABOUR LAW IN EUROPE* (2001); Davidov, *supra* note 125, at 3-4.

really believe that human subjects should receive some worker-type protections, applying existing worklaw might paradoxically fail, or at least not reach all the right subjects.

B. Shared Concerns of Subjects and Protected Workers

Because the prevailing legal tests for distinguishing between different types of workers fail to generate results that correspond reliably and precisely to those who really ought to be protected – and would likely do the same for subjects given critical questions regarding who is in control and the temporary nature of research relationships – it is necessary to look beyond these tests in both contexts. But if we acknowledge that “there are some workers (‘employees’) . . . that are in need of protection, while at the same time there are others who are capable of protecting themselves in a market environment [and] . . . in the case of employees there is a corresponding employer who can and should take responsibility for their well-being,”¹³⁶ it is insufficient to simply argue, as the bioethicists have, that research subjects are like workers and should be treated as such. Instead, it is necessary to demonstrate that they are more like those workers who *should* be protected than like those who will do fine on their own.

Leaving articulation of the perfect dividing line between protected and unprotected workers to the worklaw scholars, it seems that subjects often share not only the basic features of workers in general – at least some are engaged in commercial transactions, are paid to provide a service, and view what they do as a job – but also the very same features of workers who are unable to protect themselves and therefore deemed to need legal intervention. For example, they have divergent interests from those in positions of control, highly unequal bargaining power, strong potential for exploitation, democratic deficits, dependence, and limited alternatives. Moreover, there exists a party that can reasonably be expected to take responsibility for subjects’ well-being: research sponsors. Even if sponsors are not necessarily seeking to profit from the research, and even though the burdens placed on subjects may be justifiable in light of the ends pursued, there is nothing unfair about expecting sponsors to protect the subjects used to achieve the sponsor’s goal – or at least nothing less fair than expecting employers to protect their employees even when the work done has some socially valuable purpose and benefits that accrue to others.

1. Divergent Interests

Both workers and subjects have interests that often diverge from the interests of those in control, although this is not to say that there is no overlap. Employees and employers may share the goal of profitability, and subjects and researchers may share the mutual goal of safety. However, managers will generally extend

136 Davidov, *supra* note 125, at 13; *see also* Carlson, *supra* note 81, at 356.

only those benefits necessary to attract and retain an acceptable workforce, which is precisely why the law must sometimes intervene.

Similarly, subjects and researchers or sponsors may not necessarily be overt antagonists, but they are not always on the same side of the line. For example, more amenities offered at a Phase I research facility, higher payment to subjects, more safety tests, and the like will all lead to a more expensive trial, leaving less profit for the drug company or CRO – or perhaps of greater concern, less money available for other research. Other examples abound. Including greater detail in a consent form might help shield an institution from lawsuits or liability,¹³⁷ but make it more difficult for the average subject to understand.¹³⁸ Sponsors have little incentive to investigate or track latent effects of trial participation, although that may be important to subjects' long-term health. And even the most upstanding researcher is (or should be) focused on the population-level benefits of increased scientific knowledge;¹³⁹ in order to obtain that knowledge, subjects are inherently asked to do things that may run against their own physical interests, such as undergo tests needed only to generate data rather than to improve their care. Thus, like employees and employers, subjects and those doing the research are in many cases after different things. This is true regardless of whether the subjects are healthy or patients, paid or unpaid, or in therapeutic or nontherapeutic studies.

2. Bargaining Power, Collective Action, and Exploitation

In some cases, workers and subjects may also share the plight of poor bargaining power and collective action problems. As a matter of simple supply and demand, when there is a surplus of workers willing to do a job, competition between them can create a race to the bottom. Addressing this problem is another reason worklaw often intervenes in the market.¹⁴⁰

137 See Michelle M. Mello, David M. Studdert & Troyen A. Brennan, *The Rise of Litigation in Human Subjects Research*, 139 ANNALS INTERNAL MED. 40 (2003).

138 See Nancy E. Kass et al., *Length and Complexity of US and International HIV Consent Forms from Federal HIV Network Trials*, 26 J. GEN. INTERNAL MED. 1324 (2011) (demonstrating consent forms to be long and in excess of recommendations for how much information can be readily processed). But see Leanne Stunkel et al., *Comprehension and Informed Consent: Assessing the Effect of a Short Consent Form*, IRB: ETHICS & HUM. RES., July–Aug. 2010, at 1 (finding that comprehension was generally poor and that “neither comprehension of study information nor satisfaction with the consent process was affected by either the length or the complexity of the consent form”).

139 Franklin G. Miller & Howard Brody, *A Critique of Clinical Equipoise: Therapeutic Misconception in the Ethics of Clinical Trials*, 33 HASTINGS CENTER REP. 19, 21 (2003).

140 Davidov, *supra* note 125, at 5 (“It is often said that the basic characteristic of an employment relationship – which is also the background reason for all protective labour and employment regulations – is the inequality of bargaining power between the individual employer and the individual employee.”). However, Davidov contends that inequality of bargaining power is not necessarily helpful in determining which class of workers should be entitled to “employee” protections. *Id.* at 8.

Similarly, in the research context, there is often competition to get into studies, either for therapeutic or financial reasons.¹⁴¹ When subjects are replaceable, and particularly when their alternatives are limited, they may be unwilling to risk exclusion in order to negotiate for better terms, individually or collectively.¹⁴² Once a study has begun, subjects are in a better position considering the sunken investment in their data,¹⁴³ but even then, loss of a single subject is a relatively surmountable obstacle for those conducting the research. The real concern would be losing all the subjects, and thus all the data, and having to start from scratch. However, that requires collective action, which may be difficult to muster.

The ultimate concern raised by unequal bargaining power is that subjects and workers may find themselves in circumstances of mutually advantageous consensual exploitation. In other words, they will benefit enough to make participation worth their while, but it will nonetheless be unfair if they deserve more given the burdens they are undertaking or the value of their work.¹⁴⁴ The baseline entitlements and protections of worklaw help to address the possibility of such exploitation for workers, but many important entitlements are lacking for HSR.

3. Democratic Deficits, Dependency, and the Need for Protection

Finally, research subjects may face the two fundamental problems that worklaw scholar Guy Davidov identifies as justifying the extension of worker protections: “democratic deficits” and economic dependency.¹⁴⁵ He defines democratic deficits to mean subordination to or control by another, and economic dependency in terms of one’s ability to spread economic risks among a number of different relationships.¹⁴⁶ Thus, dependency is determined by the exclusivity of the arrangement, the proportion of income derived from a particular hiring party, and the duration of the engagement.¹⁴⁷

It seems clear that subjects often meet the subordination criterion, but dependency may be more difficult, particularly given Davidov’s narrow understanding of the term. First, some subjects may not be economically dependent on research participation at all; they may not even be paid. But they may nonetheless be dependent on a trial for other benefits, such as access to

141 Ondrusek, *supra* note 27, at 85; Abadie, *supra* note 32.

142 Elliott & Abadie, *supra* note 45, at 2317; Lemmens & Elliott, *supra* note 43, at 52; VanderWalde, *supra* note 47, at 26-27; Ondrusek, *supra* note 27, at 121.

143 ABADIE, *supra* note 27, at 57-58.

144 Elliott & Abadie, *supra* note 45, at 2317; Alan Wertheimer, *Exploitation in Clinical Research*, in *EXPLOITATION AND DEVELOPING COUNTRIES: THE ETHICS OF CLINICAL RESEARCH* 63-104 (Jennifer S. Hawkins & Ezekiel J. Emanuel eds., 2008).

145 Davidov, *supra* note 133, at 63.

146 *Id.* at 61-63.

147 *Id.* at 67-68.

certain types of medical interventions unavailable elsewhere. Second, paid subjects may not be dependent on any particular trial, but rather on participation in trials in general. Similarly, a worker may have several jobs and thus be better able to withstand the loss of any one of them, while being relatively powerless in all of them because her alternatives are no better and may be worse. This creates a certain type of dependency that is admittedly broader than that articulated by Davidov, but nonetheless seems relevant.¹⁴⁸

In a perfectly competitive market, there are enough other workers, enough other jobs, enough information, and few enough barriers that the parties will reach an ideal agreement; there is no need for legal intervention. Of course, this is often not the case,¹⁴⁹ and Davidov acknowledges that the presence or absence of asymmetric vulnerability is an important difference when it comes to allocating legal protection between different types of workers.¹⁵⁰ However, he maintains that a focus on unequal bargaining power is problematic for determining the proper application of worklaw protections because it is an empirical question and, more importantly, because some degree of inequality almost always exists in any market. Thus, Davidov proposes the dependency criterion as an alternative that may be more readily discernable,¹⁵¹ while still noting a link between the concepts: “dependency points to inequality of bargaining power; to the ineffectiveness of both voice and exit as means to protect one’s self.”¹⁵² And research subjects may have problems with both. Ultimately, a feature they often share with other workers who are and should be granted worklaw protections is lack of access to better alternatives.

C. Weaker Analogy, Same Protections

Thus, many subjects are like those workers to whom the law does and/or ought to extend protection, but this is not true across the board. Should this mean that only some subjects ought to be granted analogous new rights? Not necessarily.

First, consider subjects motivated purely by altruism, who resemble volunteer workers to whom worklaw protections usually do not extend.¹⁵³ Without the pull of either therapeutic or financial need, these subjects have adequate bargaining power to protect their own interests; if they do not like the terms offered by a study, they can easily walk or hold out for better. Thus, it would be appropriate to refrain from extending worker protections to altruistic subjects, not because they do not deserve them, but because they can achieve these terms on their own.

148 Stone, *supra* note 89, at 284.

149 Davidov, *supra* note 125, at 6.

150 *Id.* at 8.

151 *Id.*

152 Davidov, *supra* note 133, at 67.

153 Rubinstein, *supra* note 86.

But how could we practically segregate these subjects from others to whom worker protections should be extended? A rule excluding unpaid subjects would not suffice, although it would mirror the legal rule applicable to volunteers, since subjects might still be motivated by therapeutic benefits and therefore less able to bargain effectively on their own behalf.¹⁵⁴ A rule excluding unpaid subjects in nontherapeutic research also would not work to tease out the pure altruists because subjects might nonetheless be motivated by therapeutic benefits in light of the therapeutic misconception,¹⁵⁵ or perhaps strong feelings of obligation to their doctors, treating institution, or others with their same ailment. Perhaps the best rule would be to exclude unpaid, healthy subjects in nontherapeutic research, although even they may be motivated by something other than altruism, such as free medical care or other nonmonetary benefits. Moreover, altruistic subjects may be paid, they are just not (primarily) motivated by that fact. Ultimately, altruistic subjects do not need worker protections, but it is likely better to be over-inclusive here, since extending worker protections to altruists would simply grant them at least what they would demand in order to be willing to participate in the first instance, whereas excluding them could potentially sweep up some subjects who really are in need.¹⁵⁶

On the other hand, some subjects who would benefit from the extension of worker protections and could not effectively bargain for them actually seem different from protected workers in important ways. For example, unpaid subjects lack a defining characteristic of the protected worker: payment for labor. And subjects in therapeutic research may be paid, but also differ from protected workers in that the benefits they predominantly anticipate are completely outside the commercial context. Thus, it is not quite right to suggest that these subjects ought to be treated just like protected workers because they are just like protected workers. They are not. Nevertheless, payment to workers is not itself what renders them in need of protection; instead, it is dependence on that payment, which in turn reduces their bargaining power. Subjects – paid or not, seeking noncommercial benefits or not – can be similarly dependent, with similar outcomes.

Clearly, the law does not intervene every time there is a disparity in bargaining power, but it often does. Moreover, it specifically protects weak workers, albeit imprecisely, from a number of problems that also threaten weak subjects. And perhaps most importantly, the government is far from hands off in the research context. For decades, subjects have been viewed as in need of stringent regulatory protection, and intrusion in the relationship between

154 Note that a similar problem arises when volunteers actually need the volunteer opportunity for some other purpose, such as college credit for an internship, which impedes on their ability to bargain and exit. See, e.g., *id.* at 151.

155 Paul S. Appelbaum & Charles W. Lidz, *The Therapeutic Misconception*, in *THE OXFORD TEXTBOOK OF CLINICAL RESEARCH ETHICS*, *supra* note 18, at 633-44.

156 See Davidov, *supra* note 125, at 11.

researcher and subject is commonplace, as exemplified by the regulatory requirements for IRB review and approval, specific standards for acceptable consent, control over subject recruitment and compensation, and the like. Against this background, it seems that subjects should be at least as well protected as workers when they face similar problems, even when the analogy is imperfect.

IV. LEVELING SUBJECTS UP TO WORKERS

When groups that are the same in relevant respects face uneven treatment, consistency demands that the unevenness be remedied, which can occur either by leveling up or leveling down. For example, in those cases where workers are more protected than subjects – our focus here – the difference could be eliminated by granting subjects those same protections or stripping those protections from workers. However, considering that workers really seem to need the protections they are granted, the latter option of leveling down would be unwarranted and inappropriate. Moreover, considering that we have far more experience with regulation of the workplace than with regulation of human subjects research, and because many more people are traditional workers than are research subjects, existing worklaw protections are likely to be a better reflection of social norms and agreement than the absence of those protections in the research setting.¹⁵⁷ Thus, the remaining options are: (1) to level subjects up for parity with the legal status quo applicable to workers, while acknowledging that workers (and therefore subjects) should sometimes be even more protected than they currently are, or (2) to identify some relevant distinction justifying a particular difference in treatment. What follows will consider which of these approaches is warranted for some of the most pressing and obvious disparities between the two groups.

A. Payment

One of the most highlighted differences between research participation and employment – indeed, the one that got Lemmens and Elliott going – is the disparity in how payment is regulated in either case. The HSR regulations permit subject payment, but include no minimum wage requirement. They also impose no maximum, but the general regulatory requirement to protect consenting subjects from undue inducement in practice acts as a variable payment ceiling.¹⁵⁸ Moreover, although technically silent on the matter of whether payment to

¹⁵⁷ Sachs, *supra* note 7, at 75.

¹⁵⁸ 45 C.F.R. § 46.116 (2013); 21 C.F.R. § 50.20 (2013); OHRP, *Informed Consent*, *supra* note 42; OHRP, *Institutional Review Board Guidebook: Chapter III*, HHS § G (1993), http://www.hhs.gov/ohrp/archive/irb/irb_chapter3.htm [hereinafter OHRP, *IRB Guidebook*] (section on “incentives for participation”); see also Emily A. Largent et al., *Money, Coercion, and Undue Inducement: Attitudes about Payments to Research Participants*, IRB: ETHICS & HUM. RES., Jan.–Feb. 2012, at 1 (finding that IRB members are very concerned that subject payments will lead to undue inducement, which may result in limiting the payments offered).

subjects may be based on risk, the regulations' direction to avoid undue inducement is often taken to mean that risk-based payment is impermissible.¹⁵⁹ Federal regulators have recently made an effort to clarify that remuneration to subjects may indeed include compensation for risks, and that compensation may be treated as an acceptable motive for subjects agreeing to participate in research. However, IRBs are nonetheless cautioned that such remuneration should not be treated as offsetting research risks in the analysis that boards themselves are required to undertake before approving research proposals. And still, IRBs are warned to ensure that "payments are not so high that they create an 'undue influence' or offer undue inducement that could compromise a prospective subject's examination and evaluation of the risks or affect the voluntariness of his or her choices."¹⁶⁰ Ultimately, the concern is that the offer of high payment may be so irresistible to subjects that it will lead them to exercise poor judgment in accepting unnecessary, unreasonable, and excessive risks of serious harm, *i.e.*, risks that a reasonable person would not assume.¹⁶¹

Whether undue inducement ought to be a relevant consideration in the research context is a matter of significant debate,¹⁶² but what is clear is that fear of undue inducement plays *no role whatsoever* in the legal regulation of wages paid to workers. In fact, worklaw imposes no ceiling – explicit or implicit – on how much workers may be paid, although other factors may. And in theory, the market should dictate (and some laws do)¹⁶³ that risky work be better compensated, a phenomenon called the compensating wage differential.¹⁶⁴ Further, even when risky jobs are held by those with few other options for less risky work that is comparably compensated, the law does not require that their payment be restricted on that basis.¹⁶⁵ Worklaw does, however, impose a compensation floor in some cases. Workers meeting the Fair Labor Standards Act's definition of nonexempt employee are entitled to be paid the federal minimum wage of \$7.25 per hour (since 2009), and at least one and a half times

159 See Sachs, *supra* note 7, at 70.

160 OHRP, *Informed Consent*, *supra* note 42.

161 Ezekiel J. Emanuel, *Undue Inducement: Nonsense on Stilts?*, 5 AM. J. BIOETHICS 9 (2005); OHRP, *IRB Guidebook* *supra* note 158; see also Dickert & Grady, *supra* note 29, at 390 (providing a somewhat different definition of undue inducement as characterized by a choice to engage in an activity even though a subject finds it objectionable in some significant way).

162 See, e.g., Emanuel, *supra* note 161; Ruth Macklin, 'Due' and 'Undue' Inducements: On Paying Money to Research Subjects, IRB: ETHICS & HUM. RES., May 1981, at 1; McNeill, *supra* note 7, at 393; Savulescu, *supra* note 46.

163 See, e.g., 5 C.F.R. pt. 550, subpart I (2013) ("Pay for Duty Involving Physical Hardship or Hazard").

164 But see Kevin Purse, *Work-Related Fatality Risks and Neoclassical Compensating Wage Differentials*, 28 CAMBRIDGE J. ECON. 597 (2004).

165 See Miller & Wertheimer, *supra* note 22, at 91 ("Note that we do not say that people are coerced into taking jobs because they would otherwise be poor or unemployed.").

the regular rate of pay for overtime worked beyond 40 hours per week;¹⁶⁶ state and local minimum wages may be even higher.

Then again, there are circumstances in which employment law permits workers to go without payment, provided that they are either true volunteers altruistically contributing to public welfare or are otherwise receiving adequate nonmonetary compensation in the form of training or reputational benefit.¹⁶⁷ Unpaid research participation may also be acceptable under similar circumstances,¹⁶⁸ such as non-profit research conducted for the common good or other types of research participation that offer therapeutic promise directly to subjects adequate to compensate for their contribution. There are, of course, no such prerequisites for unpaid research participation under the current regulatory scheme. But even if we leave aside those circumstances in which subjects are not paid but should be, the question remains: when a decision has been made to offer payment for research participation,¹⁶⁹ is there any reason subjects should not be entitled to the same payment protections and liberties as paid workers?

One possibility is simple paternalism; we do not want subjects to exchange their health for money, so we seek to discourage them from doing so to the extent compatible with achieving sufficient enrollment. But even if one accepts paternalistic justifications for regulation as legitimate, they fail to explain the difference in treatment between subjects and other workers who also face various physical risks for pay. There may also be some initial concern regarding topics like commodification of the research subject, crowding out altruists, or transformation of what should be a gift relationship into a commercial

166 Wage & Hour Div., *Compliance Assistance – Wages and Fair Labor Standards Act (FLSA)*, U.S. DEP'T LAB., <http://www.dol.gov/whd/flsa/> (last visited Nov. 22, 2013).

167 See, e.g., Office of the Assistant Sec'y for Policy, *Fair Labor Standards Act Advisor: Volunteers*, U.S. DEP'T LAB., <http://www.dol.gov/elaws/esa/flsa/docs/volunteers.asp> (last visited Nov. 22, 2013); Wage & Hour Div., *Fact Sheet #71: Internship Programs Under the Fair Labor Standards Act* U.S. DEP'T LAB. (Apr. 2010), <http://www.dol.gov/whd/regs/compliance/whdfs71.pdf>.

168 Some argue that unpaid research is one way to avoid subject exploitation. See Trisha Phillips, *Exploitation in Payments to Research Subjects*, 25 *BIOETHICS* 209, 217-218 (2011) ("Avoiding exploitation does not require that researchers pay a fair wage; it merely requires that they do not pay less than a fair wage.").

169 On the general question of whether subject payment is itself ethically permissible, there is a tremendous body of literature considering aspects like the impact of payment on risk assessment and consent, the tension between protecting against undue inducement and avoiding exploitation, and various mechanisms of setting payment rates to avoid these problems. See ABADIE, *supra* note 27, at 6-7, 65-84; Dickert & Grady, *supra* note 29, at 386-96; Emanuel, *supra* note 161; Scott D. Halpern, *Financial Incentives for Research Participation: Empirical Questions, Available Answers and the Burden of Further Proof*, 342 *AM. J. MED. SCI.* 290 (2011); McNeill, *supra* note 7; Open Peer Commentaries, *Money for Research Participation: Does It Jeopardize Informed Consent?*, 1 *AM. J. BIOETHICS* 45, 45-68 (2001) (multiple authors); Phillips, *supra* note 168; VanderWalde & Kurzban, *supra* note 26; Martin Wilkinson & Andrew Moore, *Inducement in Research*, 11 *BIOETHICS* 373 (1997); Ondrusek, *supra* note 27, at 18-34, 188-201.

transaction.¹⁷⁰ However, so long as the regulations permit subjects to be paid beyond simple reimbursement of their expenses, it seems the ship has sailed on these concerns, and the focus should instead be on ensuring fair payment amounts. Moreover, it really should be no more worrisome to commodify a person's labor as a research subject than to commodify a person's labor in other contexts, which happens all the time. Crowding out seems also to be a red herring given that at least one of the reasons payment is currently offered is that altruistic research participation does not suffice to meet demand,¹⁷¹ not to mention the fact that altruistic subjects are always free to reject payment, if they so choose. And there seems to be no discernible reason that research participation ought to be treated as a gift when people prefer to be paid. Finally, and most importantly, it is essential to recognize that none of these concerns actually drive the *regulation* of payment to subjects, which instead is rooted exclusively in fears of undue inducement. And so the question stands: why should paid subjects and paid workers not be treated the same with regard to that payment?

1. *Payment Ceiling*

Breaking down the different aspects of payment, note that the concept of a payment ceiling might be viewed as unique among the disparities between the HSR and worklaw regimes considered herein. Elsewhere, we focus on circumstances in which subjects are decidedly less protected than workers. However, the absence of a payment ceiling in the context of most traditional work might be characterized as a freedom rather than a protection, whereas the presence of a ceiling for research participation is generally billed explicitly as a subject protection. Thus, in this context, subjects might appear to be *more* protected than workers at present, throwing a wrench into the rationale for maintaining that it is the subject side of the analogy that should be open to regulatory amendment and enhanced protections; eliminating the payment ceiling could be viewed as leveling subjects down. On the other hand, there is also an important sense in which restricting subject payment is not protective at all: payment restrictions open subjects up to the exploitative possibility of being paid too little, in which case it is the *absence* of a payment ceiling that is more protective, as in the work setting. Given the dual nature of the payment restrictions applicable to HSR – protective against undue inducement but permissive of exploitation – it is necessary to step back and confirm that worklaw provides the appropriate fixed standard when it comes to resolving payment

170 See, e.g., Tod Chambers, *Participation as Commodity, Participation as Gift*, 1 AM. J. BIOETHICS 48 (2001). For a discussion of these arguments and why they fail to prove that payment to subjects at any level is inherently problematic, see Dickert & Grady, *supra* note 29, at 391.

171 Note that the fact that research subjects can be paid places HSR in stark contrast to organ donation, where crowding out is an oft-cited argument against initiation of an organ market, albeit one with scant empirical evidence to support it. See Julia D. Mahoney, *Altruism, Markets, and Organ Procurement*, 72 LAW & CONTEMP. PROBS. 17, 24-26 (2009).

inconsistencies between subjects and workers. In other words, should we be considering a payment ceiling for workers too?

Note that a maximum wage for work is not completely unheard of. In fact, maximum wage laws existed in colonial America,¹⁷² and were also passed by some state legislatures in the South after emancipation for reasons that were decidedly anti-worker.¹⁷³ More recently, some have suggested a cap on executive compensation to help avoid adverse effects on economic conditions and financial stability.¹⁷⁴ Similarly, many professional sports leagues have adopted private pay scales and salary caps that limit how much any player may receive in order to keep more teams in the competitive range and preserve entertainment value.¹⁷⁵ And government employees (and grant recipients) are also subject to maximum pay rates to ensure judicious use of citizens' tax dollars.¹⁷⁶

In none of these examples, however, is payment limited out of fear that workers will suffer from undue inducement. In fact, when it comes to risky work, many would suggest that workers deserve to be paid substantially more, not less.¹⁷⁷ Higher pay would likely also attract a broader swath of the population to risky work, potentially allowing risks to be more evenly distributed rather than concentrating them on the very worst off. Thus, rather than introducing fresh concerns regarding undue inducement into the employment setting, it is more appropriate to treat freely paid risky work as the fixed comparator. The next step is to assess whether there is any reason to be more concerned when subjects accept risks because they want or need the money, and whether this heightened concern would justify retention of the payment ceiling in the research context alone.

A number of reasons have been suggested in the literature, all of which fail. For example, it cannot be that risks to subjects are greater, because the greatest risk – death – is also present in some jobs. And it cannot be that the risks to subjects are unreasonable or cannot be minimized, since this is specifically regulated by IRBs. Nor can it be that risks to subjects are more uncertain or

172 William P. Quigley, *Work or Starve: Regulation of the Poor in Colonial America*, 31 U.S.F. L. REV. 35, 69 (1996).

173 James Gray Pope, *Contract, Race, and Freedom in the Constitutional Law of "Involuntary Servitude,"* 119 YALE L.J. 1474, 1534-1535 (2010).

174 Sarah B. Patterson, Note, *Protecting Your Rights, But Not Your Paycheck: How Executive Compensation Regulation Passes Constitutional Muster*, 113 W. VA. L. REV. 931 (2011); Sam Pizzigati, *The Corporate Pay Gap: Do We Need a Maximum Wage?*, PERSP. ON WORK, Summer 2009, http://www.lera.uiuc.edu/pubs/perspectives/CompArticles/POW_13.1--Pizzigati.pdf.

175 See John Clayton, *What New CBA Means in Football Terms*, ESPN (July 25, 2011), http://espn.go.com/nfl/story/_id/6790759/what-new-nfl-cba-means-football-terms.

176 See, e.g., 5 C.F.R. § 531.221 (2013) ("Maximum Payable Rate Rule"); *Frequently Asked Questions: NIH Salary Cap in FY2012*, NAT'L INSTS. HEALTH, http://grants.nih.gov/grants/policy/fy2012_salary_cap_faqs.htm (last updated Feb. 14, 2012).

177 Some argue this is true for research as well. See, e.g., Eleri Jones & Kathleen Liddell, *Should Healthy Volunteers in Clinical Trials Be Paid According to Risk? Yes*, 340 BRIT. MED. J. 130 (2010); Jerry Menikoff, *Just Compensation: Paying Research Subjects Relative to the Risks They Bear*, 1 AM. J. BIOETHICS 56 (2001); Ondrusek, *supra* note 27, at 29, 192.

unknown.¹⁷⁸ First, it is unclear why a range of possible risk from minor to severe should be more worrisome than a known risk concentrated at the higher end of the spectrum, which is true of some jobs. Second, just because research involves uncertainty does not necessarily mean that sensible predictions are impossible.¹⁷⁹ Third, “it is probably false that research risks are, in general, poorly understood compared to the risks undertaken by employees and volunteers. In fact, some employees and volunteers, such as test pilots, take risks every day without knowing their extent.”¹⁸⁰ And finally, to the extent that the inherent uncertainty of research risk is not deemed to invalidate subject consent, it does not seem to exacerbate the potential for undue inducement.

Perhaps the difference is that subjects are taking risks for the benefit of others, so we should be especially concerned that those risks are undertaken voluntarily, but that rationale fails as well. Firefighters also take risks for the benefit of others, and paid subjects are often taking risks for their own financial benefit. There may be a concern about therapeutic misconception in research that is not present in the context of work,¹⁸¹ but if anything, payment could help make clear that research is different from clinical care.¹⁸² And for both healthy and patient-subjects, payment amount may be an important indication of risk level.¹⁸³ Some worry that payment might blind subjects to potential risks, whereas workplace risks may be more obvious, but empirical evidence on this is scant,¹⁸⁴ and if it is a legitimate concern, the appropriate solution is to improve research consent before restricting payment.¹⁸⁵ Some also maintain that it is disrespectful to offer people money to entice them to overcome deeply held objections based on their values, desires, or fears.¹⁸⁶ However, it happens all the time – just consider the number of law school graduates who head straight for a high-paying firm job knowing that the hours will make them unhappy or that they will pursue goals for their clients that they personally disagree with. Ultimately, none of these reasons convincingly suggest that avoidance of undue inducement is a reason to limit payment to subjects but not to workers, or for that matter, to reject

178 McNeill, *supra* note 7, at 391.

179 Timothy Wilkinson, *Assessing the Case for the Regulation of Research*, 10 AM. J. BIOETHICS 63, 64 (2010).

180 Sachs, *supra* note 7, at 74.

181 *Id.* at 74-75.

182 Dickert & Grady, *supra* note 29, at 392 (also noting that payment to patient-subjects could “depersonalize the exchange, making it easier for patients to refuse and putting them on more equal bargaining terms with investigators”).

183 See Cynthia E. Cryder et al., *Informative Inducement: Study Payment as a Signal of Risk*, 70 SOC. SCI. & MED. 455 (2010).

184 See Scott D. Halpern et al., *Empirical Assessment of Whether Moderate Payments Are Undue or Unjust Inducements for Participation in Clinical Trials*, 164 ARCHIVES INTERNAL MED. 801 (2004) (finding no evidence that commonly used payment levels represent undue inducements).

185 Dickert & Grady, *supra* note 29, at 389.

186 *Id.*

risk-based payments to subjects when such payments are expected and encouraged for other work.

There are, however, a few possible concerns beyond undue inducement that have been suggested as rendering unrestricted payment particularly worrisome in the research setting. First, some subjects motivated by money may lie or withhold information in order to enroll or stay in a trial,¹⁸⁷ and/or enroll in multiple trials, leading to both methodological and ethical concerns. For example, the validity of trial data may be jeopardized if subjects fail to disclose prior or concurrent study participation, evade exclusion criteria, violate study requirements, or fail to report side effects. These behaviors might also hurt subjects themselves.¹⁸⁸

But does this really set research apart? Workers may lie about their qualifications too, in ways that put both themselves and their employers' output in jeopardy, and they may be enticed to do so by money. And their lies are in many cases objectively detectable; a manager can ask for training certificates, speak with previous employers, test competence, carefully oversee the work, and the like. Similarly, although US regulations do not currently address the issue of repeat or simultaneous participation, regulatory bodies, sponsors, and researchers could implement national subject registries to track participants,¹⁸⁹ impose and enforce mandatory wash-out periods between trials, institute lifetime enrollment caps, utilize more extensive screening before enrollment, and increase use of physical testing rather than relying on qualitative subject feedback whenever possible.¹⁹⁰ Unless these solutions were unsuccessful, and unless empirical evidence suggests that scientific integrity and subject safety are indeed being harmed by the offer of unrestricted payment for participation, the fact that paid subjects might be more likely to lie than those who are unpaid cannot justify a limit on compensation to subjects but not for other jobs.

What about concern that without an upper limit on payment, some important research will likely be unable to compete for subjects?¹⁹¹ This would be regrettable, but this problem is not unique to research either. In all sorts of jobs, the public sector and non-profits must compete for workers with the private sector and profit-driven companies. They often do so in non-monetary ways, for example, by emphasizing civic duties and the importance of helping others. Even when one company's mission might be clearly more desirable (from a moral perspective) than another's, we do not limit the payment that can be offered by

187 *Id.* at 390; Carl L. Tishler & Suzanne Bartholomae, *Repeat Participation Among Normal Healthy Research Volunteers: Professional Guinea Pigs in Clinical Trials?*, 46 *PERSP. BIOL. & MED.* 508, 512-13 (2003); Ondrusek, *supra* note 27, at 42, 90, 103, 143. Note, however, that this could be an issue for any subject motivated for reasons other than altruism.

188 ABADIE, *supra* note 27, at 154-55; Elliott, *supra* note 30; Elliott & Abadie, *supra* note 45, at 2317; Tishler & Bartholomae, *supra* note 187, at 511, 512, 514-15; Wilkinson & Moore, *supra* note 169, at 388; Ondrusek, *supra* note 27, at 42, 90, 103, 143.

189 See Resnick & Koski, *supra* note 38; Tishler & Bartholomae, *supra* note 187, at 517.

190 Tishler & Bartholomae, *supra* note 187, at 513.

191 Dickert & Grady, *supra* note 51, at 201; Trisha B. Phillips, *A Living Wage for Research Subjects*, 39 *J.L. MED. & ETHICS* 243, 243 (2011).

either. As a result, some socially valuable projects simply cannot flourish. Unless clinical research is somehow different from these other important projects, which may also aim to save lives or otherwise improve the world, competition for subjects is not a reason to limit payment for participation. Note that the same rationale applies to the concern that high payment to subjects will drive up the cost of doing research and take resources from other important projects.

Ultimately, regardless of whether there is reason for genuine concern about undue inducement of research subjects and/or other potential negative effects of unrestricted or risk-based payment, there seem to be only two possible reasons to treat research differently from other types of work in this regard, and both are somewhat hypothetical. If high payments would in fact damage scientific integrity, or if some essential research that is deemed more important than other endeavors cannot compete for subjects, then it would be possible to justify regulatory restrictions on subject payment even when such restrictions are not present for other types of work. And of course, some types of payment restriction for research participation would be completely consistent with a work regime, such as a private “salary cap” negotiated between a group of sponsors and subjects,¹⁹² or a regulatory cap for subjects in government-funded research. But these circumstances would be exceptional, whereas restricted payment is now the rule for HSR. Thus, we have our first case in which the subject-worker analogy calls for a substantial rule change.¹⁹³

2. *Minimum Wage*

Removing payment restrictions for research subjects will eliminate one excuse for payments that are too low. But, focusing again on consistency and the need to treat like cases alike, should subjects also be guaranteed the minimum wage? The reasons that supported unrestricted subject payment as discussed above seem to suggest that here too the answer is yes. However, there are some relevant differences in purpose that indicate the minimum wage need not be extended to all paid subjects.

States began to introduce minimum wage laws in the early twentieth century out of concern that many workers who were unable to effectively bargain with their employers were receiving a wage below that necessary to provide an

192 A private salary cap would be permissible only if agreed upon in collaboration with subjects through a labor agreement or if Congress adopted some exemption to antitrust laws for the research context. See *Labor and Collective Bargaining*, AM. ANTITRUST INST., <http://www.antitrustinstitute.org/content/labor-collective-bargaining> (last visited Nov. 22, 2013).

193 Others have also argued that there should be no upper limit on subject payment, but not necessarily for all the same reasons articulated here and not necessarily for all subjects. See, e.g., Ackerman, *supra* note 28, at 1, 3; Anderson & Weijer, *Wage Earner*, *supra* note 57, at 375 n.8; Elliott & Abadie, *supra* note 45, at 2317; Lemmens & Elliott, *supra* note 43, at 53; David B. Resnick, *Research Participation and Financial Inducements*, 1 AM. J. BIOETHICS 54, 55 (2001); Shamoo & Resnick, *supra* note 34, at W10.

adequate standard of living. The federal minimum wage was initially stimulated by a desire to raise purchasing power and boost the economy, but the rationale eventually expanded to include poverty reduction goals as well.¹⁹⁴ Among other things, Congress noted that “labor conditions detrimental to the maintenance of the minimum standard of living necessary for health, efficiency, and general well-being of workers” burden commerce, constitute an unfair method of competition, and lead to labor disputes.¹⁹⁵

Importantly, these rationales for the minimum wage seem to apply only to those research subjects who participate in research instead of or to complement other work. For these subjects, the amount paid really does influence their standard of living and purchasing power. In contrast, the rationales for a minimum wage do not fit subjects who participate in research primarily for altruistic or therapeutic reasons, even though they may also be paid and even when payment might have pushed them over the tipping point in agreeing to enroll.¹⁹⁶ The difference is really whether subjects experience payment as a bonus or as a wage. This could be operationalized, albeit somewhat imperfectly, by extending the applicable minimum wage only to paid healthy subjects participating in nontherapeutic research. Not everyone in this category may actually need to be guaranteed the minimum wage, since at least some research currently pays more and at least some subjects will select alternative opportunities when the pay offered is insufficient. However, the same is true for other work where the minimum wage is nonetheless extended because some workers do need it – if Mark Zuckerberg flipped burgers for McDonald’s, he would be entitled to the same payment protections as everyone else.

That being said, there would be a few complexities associated with extending the minimum wage to research subjects. For example, some of those who argue that subjects currently get paid too little actually break down the subjects’ compensation into an hourly rate and call for overtime pay even though subjects are not actively engaged in participation during the entire period of their enrollment.¹⁹⁷ Even in a confinement study, subjects are usually free for several hours a day to pursue leisure activities of their own choosing, and in any study, each hour of time spent as an enrolled subject is not worthy of equal compensation. Thus, the question is whether subjects are closer to the security guard¹⁹⁸ who gets paid for an entire shift even if he does no more than read a book, or to the on-call employee who may not be entitled to the minimum wage

194 John Foley, Note, *Questioning the Merits of Federal Minimum Wage Legislation*, 5 GEO. J.L. & PUB. POL’Y 679, 681-82 (2007); David Neumark & William Wascher, *Minimum Wages and Low-Wage Workers: How Well Does Reality Match the Rhetoric?*, 92 MINN. L. REV. 1296, 1298-1303 (2008).

195 29 U.S.C. § 202(a) (2006).

196 This distinction resolves potential uncertainty as to how to appropriately compensate subjects monetarily when they are getting nonfinancial benefits from participation.

197 Anderson & Weijer, *Entrepreneur*, *supra* note 57, at 68.

198 See ABADIE, *supra* note 27, at 2-3.

or overtime for hours spent on call.¹⁹⁹ A further question is whether subjects are confined for their own benefit so that they may be cared for in case of adverse events, or for the benefit of the study to ensure compliance.²⁰⁰

Ultimately, the standard hourly minimum wage may be a poor fit for HSR, given that subjects can be variably involved in active study visits, passive observation periods, overnight confinement, and/or tasks carried out while at home. Nonetheless, some regulatory minimum ought to exist for certain types of subjects, and compensation based on a minimum rate per procedure, per day of confinement, or per study visit may be most parsimonious. Again, for parity with protected workers, the goal is to ensure that paid healthy subjects in nontherapeutic research are paid an amount similar to what they could expect in other minimum wage jobs given the same level of commitment and exertion. Note, however, that this will not necessarily protect subjects against exploitatively low payment,²⁰¹ just as traditional workers may not be protected. This is because fair payment based on contribution, time, inconvenience, risk, discomfort, and other burdens may actually be higher than the minimum wage, which takes none of these factors into account.²⁰² In other words, the minimum wage is often too low for workers or subjects. Moreover, considering that many subjects are already getting paid amounts that make it worthwhile to participate in research instead of taking other jobs,²⁰³ it is likely that removing the payment ceiling will be the far more important change.

3. Unemployment

A final factor relevant to payment that ought to be briefly addressed (especially given that research subjects have in fact sought eligibility for it) is unemployment compensation. Employees are generally eligible for

199 Most courts have held that when an on-call employee is free to engage in personal activities, such as watching television or visiting with friends, the time spent on call is not compensable under FLSA, even when he has considerable restrictions placed on his geographic mobility and activities. However, some courts have found on-call time to be compensable when these restrictions are extreme and the employee is frequently called in to work. Stone, *supra* note 89, at 258. On the other hand, at least one court has held that even when the employee must remain at the work site when on call, his time is not compensable when he is free to “free to sleep, eat, watch television, watch VCR movies, play ping-pong or cards, read, listen to music” *Rousseau v. Teledyne Movable Offshore, Inc.*, 805 F.2d 1245, 1248 (5th Cir. 1986).

200 See, e.g., *Armour & Co. v. Wantock*, 323 U.S. 126, 132 (1944) (holding that waiting time is compensable under FLSA if it is “primarily for the benefit of the employer and his business” (quoting *Tennessee Coal, Iron & R. Co. v. Muscoda Local*, 321 U.S. 590, 597-98)).

201 Phillips, *supra* note 168, at 210-12. There is also a sense in which subjects might be exploited by payments that are too high if that would cause undue inducement. See VanderWalde & Kurzban, *supra* note 26, at 552.

202 See Phillips, *supra* note 191.

203 See Jessica Latterman & Jon F. Merz, *How Much Are Subjects Paid to Participate in Research?*, 1 AM. J. BIOETHICS 45 (2001) (finding that average subject payments exceeded the minimum wage).

unemployment when they become involuntarily unemployed.²⁰⁴ However, states differ with regard to whether an employee will be considered to have quit voluntarily when he knowingly takes a temporary job that ends as planned.²⁰⁵ Thus, failure to extend benefits to subjects may not actually be inconsistent with the treatment of other workers, at least not in all jurisdictions, given the temporary nature of research participation. But even if we assume that temporary workers should be covered, there is good reason that research subjects should not.

One of the goals of unemployment compensation is to stabilize employment, which it achieves through a stick: an experience rating that requires employers to pay an additional tax for its former employees receiving unemployment benefits.²⁰⁶ This does not fit the research context, however, because the work offered in any given study will never be permanent (even if a given subject is a “professional”), leaving nothing to incentivize through the experience rating. Moreover, it does not seem fair to expect a party to subsidize unemployment when it could not have helped a subject avoid it.²⁰⁷ Even if subjects would benefit from such protection, there is no one but the government that could be appropriately asked to pay, and other aspects of the social safety net seem better suited to help smooth the transition for those who rely on research participation for money.²⁰⁸

B. Care and Compensation for Injury

Moving on from payment for participation, another key area in which the HSR regulations seem to offer less protection than worklaw has to do with what

204 STAFF OF H. COMM. ON WAYS & MEANS, 112TH CONG., 2012 GREEN BOOK: BACKGROUND MATERIAL AND DATA ON THE PROGRAMS WITHIN THE JURISDICTION OF THE COMMITTEE ON WAYS AND MEANS, at ch. 4 (Comm. Print 2012), <http://greenbook.waysandmeans.house.gov/2012-green-book/chapter-4-unemployment-insurance/introduction-and-overview> (subsection on “Unemployment Insurance Introduction and Overview”).

205 Sachin S. Pandya, *Retrofitting Unemployment Insurance to Cover Temporary Workers*, 17 YALE L. & POL’Y REV. 907, 919-23 (1999); Stone, *supra* note 89, at 264-66.

206 Gillian Lester, *Unemployment Insurance and Wealth Redistribution*, 49 UCLA L. REV. 335, 344-45 (2001); Pandya, *supra* note 205, at 925.

207 A potential analogy here are the exceptions found in some state laws that permit seasonal employers to avoid certain unemployment insurance requirements and that disqualify seasonal employees from receiving benefits for unemployment outside the normal operating season. See generally Rex Williams, *Seasonal Unemployment Compensation: Insurance of a Known and Certain Loss*, 4 SAN JOAQUIN AGRIC. L. REV. 75 (1994).

208 Another problem is that those receiving unemployment benefits must accept “suitable” offers of work, and similar work to that lost is generally deemed suitable. Pandya, *supra* note 205, at 923-24. Thus, unemployed subjects might be compelled to enroll in other studies for which they are qualified or lose their entitlement to compensation, in violation of the voluntariness requirement imposed by the research regulations. 45 C.F.R. § 46.116(a)(8) (2013); 21 C.F.R. § 50.25(a)(8) (2013). The same is true whenever someone takes a job simply because the fact that it was offered renders them ineligible for future unemployment benefits, although there may be reason to believe that voluntariness is more important in research than in other types of work.

happens in the event of injury. On the one hand, with the exception of limited policies adopted by a handful of federal agencies,²⁰⁹ the regulations do not require that any special provisions be made for subjects harmed via research participation.²¹⁰ On the other, injured subjects are free to pursue usual legal remedies, and the regulations preclude informed consent materials from including any exculpatory language.²¹¹ However, for all litigants, the tort system is “time-consuming, adversarial, expensive, and has a tendency to under-compensate most . . . [of those injured] while over-compensating a select few.”²¹² Unfortunately, it is even more problematic for injured research subjects, who are likely to have difficulty showing that any duty to them was breached, that the research caused their injury,²¹³ that they did not assume the risk through informed consent, and most importantly, that their injury was anyone’s fault, since even perfectly conducted research can result in harm.²¹⁴ Moreover, several classes of research subjects, particularly those in federally conducted and international research, are prevented from receiving compensation altogether as a result of statutory and procedural barriers to tort litigation.²¹⁵

In contrast, most employers are legally responsible under various workers’ compensation statutes for guaranteeing payment of benefits to covered employees²¹⁶ who sustain injuries (including illness and death) that “arise out of”

209 See *Moral Science*, *supra* note 17, at 65-66 (and accompanying notes), 184-85 (describing limited provisions to provide free care to subjects injured in studies sponsored by entities such as the Department of Defense, Department of Veterans Affairs, NIH Clinical Center, Environmental Protection Agency, and NASA).

210 Cf. 45 C.F.R. § 46.116(a)(6); 21 C.F.R. § 50.25(a)(6) (requiring that informed consent include an explanation of whether any compensation is provided in the event of injury).

211 45 C.F.R. § 46.116; 21 C.F.R. § 50.20; OHRP & FDA, *Guidance on Exculpatory Language in Informed Consent*, HHS (Aug. 19, 2011), <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM271036.pdf> (draft guidance).

212 Elizabeth R. Pike, *Recovering from Research: A No-Fault Proposal to Compensate Injured Research Participants*, 38 AM. J.L. & MED. 7, 23 (2012). See also Wendy K. Mariner, *Compensation for Research Injuries*, in 2 WOMEN AND HEALTH RESEARCH: ETHICAL AND LEGAL ISSUES OF INCLUDING WOMEN IN CLINICAL STUDIES, 113, 121 (Anna C. Mastroianni et al. eds., 1994); Resnick, *supra* note 31, at 283.

213 INST. OF MED. (IOM), *RESPONSIBLE RESEARCH: A SYSTEMS APPROACH TO PROTECTING RESEARCH PARTICIPANTS* 193 (Daniel D. Federman et al. eds., 2002); Robert Steinbrook, *Compensation for Injured Research Subjects*, 354 NEW ENG. J. MED. 1871, 1872 (2006).

214 See IOM, *supra* note 213, at 188; Pike, *supra* note 211, at 23-24, 26-29; Larry D. Scott, *Research-Related Injury: Problems and Solutions*, 31 J.L. MED. & ETHICS 419, 423 (2003); VanderWalde and Kurzban, *supra* note 26, at 546.

215 See Pike, *supra* note 212, at 29-38 (referring to sovereign immunity, the Federal Tort Claims Act and its discretionary function exception, the Alien Tort Statute, and *forum non conveniens*). This problem is solidified by the Supreme Court’s decision in *Kiobel v. Royal Dutch Petroleum*, 133 S.Ct. 1659 (2013) (holding that the presumption against extraterritoriality constrains courts exercising their power under the Alien Tort Statute).

216 This includes the vast majority of employers and employees in traditional employer-employee relationships. There are, however, some statutory exceptions in both directions. For

and “in the course of their employment,” without regard to the fault of either party. In exchange, employers are immunized from tort suits for negligence in causing or contributing to the injury.²¹⁷

Before this system was in place, injured workers were seldom compensated, for a variety of reasons ranging from failure to bring suit to difficulty overcoming employers’ defenses.²¹⁸ But as industrial accidents were on the rise around the turn of the twentieth century, and it grew clear that workers may not be adequately compensated for the risks of occupational injury through their wages, the failures of tort law remedies became politically and socially unacceptable.²¹⁹ Workers’ compensation laws were adopted nationwide to provide injured workers a less expensive mechanism of swift, certain compensation by moving the system out of court and eliminating any requirement to prove fault.²²⁰ Workers are not necessarily made whole and there is no compensation for pain and suffering or other noneconomic damages, nor any mechanism for punitive damages, which benefits employers. But workers’ medical care is fully covered, as is some significant fraction of wage loss; in the event of death, survivors receive income and burial benefits.²²¹ In addition, since employers pay all benefits, either directly or through insurance, the system forces them to internalize the cost of injuries incidental to their business.²²² Unfortunately, these goals are not always perfectly achieved, particularly since disputes can remain regarding an injury’s job-relatedness and the extent of disability, necessitating litigation.²²³

It may not be flawless, but the fact remains that on the whole, injured workers are more protected by the law than injured subjects, even if they recover less than they might if successful in court. Moreover, the same goals and problems driving the workers’ compensation system seem similarly applicable to the research context. In fact, a “series of national advisory committees convened

example, employers with very few employees may not be covered, and non-employee workers such as volunteers may be. See *infra*, Part IV.B.4.

217 See *Wash. Metro. Area Transit Auth. v. Johnson*, 467 U.S. 925, 931 (1984) (describing the workers’ compensation *quid pro quo*); Ellen R. Peirce & Terry Morehead Dworkin, *Workers’ Compensation and Occupational Disease: A Return to Original Intent*, 67 OR. L. REV. 649, 653 (1988).

218 Joan T.A. Gabel et al., *The New Relationship Between Injured Worker and Employer: An Opportunity for Restructuring the System*, 35 AM. BUS. L.J. 403, 405 (1998).

219 Peirce & Dworkin, *supra* note 217, at 652, 655.

220 *Id.* at 652-53.

221 STAFF OF H. COMM. ON WAYS & MEANS, 110TH CONG., 2008 GREEN BOOK: BACKGROUND MATERIAL AND DATA ON THE PROGRAMS WITHIN THE JURISDICTION OF THE COMMITTEE ON WAYS AND MEANS, at sec. 15 (Comm. Print 2008), <http://democrats.waysandmeans.house.gov/sites/democrats.waysandmeans.house.gov/files/documents/wcss8.pdf> (section on “Workers’ Compensation”).

222 Peirce & Dworkin, *supra* note 217, at 653-54.

223 See, e.g., Gwen Forte, *Rethinking America’s Approach to Workplace Safety: A Model for Advancing Safety Issues in the Chemical Industry*, 53 CLEV. ST. L. REV. 513, 522-23 (2005); Mariner, *supra* note 212, at 122; Pike, *supra* note 212, at 46-47, 48, 49.

to consider the obligations owed in the event of research-related injury have concluded repeatedly that injured research participants are entitled to compensation for their injuries, that the tort system provides inadequate remedy, and that the United States should consider some form of no-fault compensation.”²²⁴ Indeed, unlike payment for participation, which is the source of substantial controversy, the contested question regarding care and compensation for injured subjects is not so much about desert as about scope and logistics.

Most commenters agree that there is an obligation to ensure that subjects do not individually bear the costs of medical care required to treat harms directly resulting from their research participation, without regard to fault.²²⁵ There is less agreement, however, as to whether there is any obligation to compensate subjects for economic and noneconomic harms beyond the costs of care.²²⁶ And there is even less agreement as to whether regulatory intervention is needed to ensure that obligations to injured subjects are satisfied, and if so, what the ideal intervention would look like.²²⁷ Some have suggested the workers’ compensation model,²²⁸ and since the question here is whether there is any compelling reason to treat subjects differently from other protected workers, workers’ compensation will be our focus.

224 Pike, *supra* note 212, at 10.

225 See, e.g., IOM, *supra* note 213, at 193; Hazel Beh, *Compensation for Research Injuries*, IRB: ETHICS & HUM. RES., May–June 2005, at 11, 12; Mariner, *supra* note 212, at 117; Pike, *supra* note 212; Resnick, *supra* note 31, at 286; *Ethical and Policy Issues in Research Involving Human Participants*, NAT’L BIOETHICS ADVISORY COMM’N (NBAC) 123 (2001), <http://bioethics.georgetown.edu/nbac/human/overvol1.pdf>; *International Ethical Guidelines for Biomedical Research*, COUNCIL FOR INT’L ORGS. OF MED. SCIS. (CIOMS) 78-79 (2002), http://www.cioms.ch/publications/layout_guide2002.pdf; *Moral Science*, *supra* note 17, at 62 and app. IV; World Med. Ass’n, *World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects* (Oct. 19, 2013), <http://jama.jamanetwork.com/article.aspx?articleID=1760318>.

226 For arguments in favor of compensation of economic loss beyond the costs of care, see IOM, *supra* note 213, at 193; Scott, *supra* note 214, at 424; CIOMS, *supra* note 225, at 78. . See also Ministry of Health and Family Welfare (Dep’t of Health) Notification, 2013, Gazette of India, pt. II, sec. 3(i) (Jan. 30, 2013), <http://www.elsevierbi.com/~media/Supporting%20Documents/Pharmasia%20News/2013/February/Clinical%20Trials%20Compensation%20Guidelines.pdf> (going to extreme lengths to demand care and compensation for injured subjects).

227 See, e.g., LEWIN GROUP, FINAL REPORT: CARE/COMPENSATION FOR INJURIES IN CLINICAL RESEARCH 21 (2005); *Moral Science*, *supra* note 17, at 64-70; Steinbrook, *supra* note 213, at 1873.

228 Bernard R. Adams & Marilyn Shea-Stonum, *Toward a Theory of Control of Medical Experimentation with Human Subjects: The Role of Compensation*, 25 CASE W. RES. L. REV. 604, 637-38 (1975); Beh, *supra* note 225, at 13; Elliott & Abadie, *supra* note 45, at 2317; Scott, *supra* note 214, at 424; Irving Ladimer, *Clinical Research Insurance*, 16 J. CHRONIC DISEASES 1229, 1233 (1963); LEWIN GROUP, *supra* note 227, at 19; Resnick, *supra* note 31, at 283-85; David B. Resnick, *Liability for Institutional Review Boards: From Regulation to Litigation*, 25 J. LEG. MED. 131, 182-83 (2004). Other approaches have also been suggested. See *Moral Science*, *supra* note 17, at 64; Pike, *supra* note 212, at 47-53.

So, is there any such reason for differential protection? The Presidential Commission for the Study of Bioethical Issues, the most recent national body to take up this question, implies that the answer is no, at least as an ethical matter. Although it refused to adopt the idea of subject as wage-earner,²²⁹ the Commission noted that if subjects are employees in a dangerous job, it would be unjust to exclude them from a form of workers' compensation available to employees in other industries.²³⁰ The Commission did not, however, go so far as to endorse the workers' compensation model for HSR, or any change at all to the status quo. Instead, it argued that before any compensation scheme can be implemented in this realm there are questions regarding "the scope of any possible coverage, the delineation of qualified harms, mechanisms for determination of causation and qualification, relation to the tort system, the need for any special public or private insurance, and how the current nonsystematic approach to this issue functions in practice."²³¹ Each of these issues is addressed in the sections below, but none provides the sort of relevant difference that can justify the disparity between the legal protections offered to injured workers and the lack of such protections for injured subjects.

1. Scope, Need, and Burden

With regard to scope of coverage, workers' compensation goes beyond what is currently agreed on (albeit not mandated) for injured subjects, but it does not go too far. This is because the same arguments that support protecting subjects from shouldering medical costs on their own seem to support protecting them and their dependents from fully bearing the weight of lost earnings, just as workers' compensation does. First, even though subjects may benefit financially or therapeutically from research, these benefits may not fully compensate for the risks they face, and even if they do, society and others also reap the benefits and should not be allowed to free-ride. Moreover, financial losses of all types are among risks to subjects that can be minimized, and beneficence and non-maleficence support a system of compensation. Finally, recruitment will potentially benefit if subjects know they will not be left to face financial risks completely on their own.²³² Thus, the scope of workers' compensation – coverage for both medical costs and a portion of lost wages – seems to be the

229 *Moral Science*, *supra* note 17, at 119 n.103.

230 *Id.* at 61.

231 *Id.* at 62. Note that the Commission's charge was specific to federally-funded research.

232 Beh, *supra* note 225, at 12; *Moral Science*, *supra* note 17, at 119 n.103, Pike, *supra* note 212, at 19-20, 56; Resnick, *supra* note 31, at 282; *see also* VanderWalde & Kurzban, *supra* note 26, at 545 (noting that the arguments in support of compensating subjects for research-related injury include encouraging research participation, relieving social discontent, fulfilling moral obligations to subjects, ensuring a just social distribution of resources, and incentivizing researchers to be careful with risk/benefit analyses).

very minimum level of appropriate compensation for injuries in the HSR setting.²³³

There are, however, two possible, mutually exclusive reasons not to implement a similar legal compensation system for injured subjects: (1) such a system is unnecessary because the needs of injured subjects are already being adequately satisfied through a patchwork of alternate mechanisms; or (2) such a system would be unduly burdensome on the research enterprise.²³⁴ The first, if true, would seem to be a relevant difference justifying stronger (or more formal) protections for injured workers than injured subjects, given that workers' needs were not adequately satisfied in the absence of legally mandated compensation. On the other hand, the second, even if true, would require some additional argument to explain why the burdens imposed on research by such a compensation scheme would be more problematic than the burdens imposed on other endeavors for which workers' compensation is required. However, since these are at least partly empirical claims, it is best to start with the empirical data.

First, it appears likely that injured subjects are *not* in fact being adequately compensated in the absence of a formal compensation system. More data are needed, but a 2005 study of over 100 academic medical center policies concluded that a subject's own health insurance serves as the "primary vehicle" for covering the cost of research-related injuries.²³⁵ Of course, not everyone has health insurance, and even if the Affordable Care Act and state initiatives are successful in achieving more universal coverage, policies vary in their inclusions and exclusions, copays, deductibles, and limits. More importantly, even if all health insurance covered clinical trial injuries,²³⁶ two problems would remain: injured subjects would still be paying for their own care via copays and other fees, and more importantly, it would remain *health* insurance, which of course does not cover other economic damages an injured subject may incur. Yet the study found that no institution or sponsor was offering to compensate injured subjects for lost wages or pain and suffering.²³⁷ And with regard to medical care, only 16% of the policies prospectively indicated a plan to provide free care or treatment for

233 One substantial concern is that requiring compensation for lost wages will drive the exclusion of high wage subjects from research and increase the recruitment of low wage subjects, with attendant justice issues. Dickert & Grady, *supra* note 51. This problem may be mitigated by capping wage recovery, precisely as workers' compensation does, although additional intervention may also be necessary.

234 *Moral Science*, *supra* note 17, at 67-68.

235 LEWIN GROUP, *supra* note 227, at ES-3. The policies in question were generally sample informed consent forms and other information available on the web sites of major medical centers. *Id.* at 4.

236 The Patient Protection and Affordable Care Act requires health insurers to pay for routine costs of care delivered in clinical trials, but it is unclear whether that includes care for study-related injury. Carmen Phillips, *Insurance Coverage Expanding for Cancer Clinical Trials*, NCI CANCER BULL. (May 18, 2010), <http://www.cancer.gov/ncicancerbulletin/051810/page5>.

237 LEWIN GROUP, *supra* note 227.

injured subjects, with an additional 10% billing a subject's insurance first, but providing free care or treatment to those without coverage.²³⁸

A similar study found that of medical schools with template language for industry-sponsored research available on their websites, 61% declare that the industry sponsor would pay the medical expenses of research-related injury. In contrast, when there is no industry sponsor, only 22% of the schools offer some form of financial support, and half of those limit coverage to emergency medical care. In addition, 72% of medical school consent forms specifically rule out the possibility of additional monetary compensation, as distinguished from free or reimbursed care, and no schools explicitly offered such compensation.²³⁹

That being said, these policies – which are themselves only a limited sample – may not necessarily reflect what happens after an injury actually occurs, when sites may be more generous than their policies dictate.²⁴⁰ And representatives of the pharmaceutical industry have recently asserted that most industry-based clinical research sponsors voluntarily or contractually agree to carry insurance to compensate individuals injured in trials, although this is not required by law.²⁴¹ Thus, on the most generous analysis, it may be the case that at least some injured subjects are in fact being adequately cared for and/or compensated for their care without having to reach into their own pockets, albeit on an ad hoc basis. Nonetheless, combining the available data with what we know about the obstacles an injured subject would face in court, the same is almost certainly not true for other economic costs of research-related injury. Accordingly, the lack of necessity argument likely fails as an empirical matter, pending further data.

However, it is important to recognize that having already established as the starting point that differential protection of subjects and workers requires justification, the onus should fall on those who claim that injured subjects do not need a workers' compensation system to demonstrate that to be the case. In other words, in the absence of complete data, the default should be to extend a workers' compensation system to injured subjects, rather than waiting until the data indicates such a system is needed. Similarly, the fact that there is little systematic and current information about the severity, frequency, and type of injuries that subjects experience and their costs²⁴² does not itself provide a reason to reject a compensation system for injured subjects until affordability has been established.

238 Steinbrook, *supra* note 213, at 1872.

239 Michael K. Paasche-Orlow & Frederick L. Brancati, *Assessment of Medical School Institutional Review Board Policies Regarding Compensation of Subjects for Research-Related Injury*, 118 AM. J. MED. 175, 177 (2005). Medical schools' IRB websites were examined for suggested text for informed consent documentation, which was then surveyed for text related to injury and compensation.

240 LEWIN GROUP, *supra* note 227; Steinbrook, *supra* note 213, at 1873.

241 *Moral Science*, *supra* note 17, at 66.

242 IOM, *supra* note 213, at vii-viii, 191-192; LEWIN GROUP, *supra* note 227, at ES-1, 2; Mariner, *supra* note 212, at 118; Scott, *supra* note 214, at 419-20; Steinbrook, *supra* note 213, at 1872.

But even if there were evidence that such a system would be very costly, hindering research more than the prospect of litigation does at present—which there is not²⁴³—that would still not necessarily suffice to justify the status quo, as some have claimed.²⁴⁴ This is because although burden is not completely extraneous to the workers' compensation analysis, it is only minimally relevant. Only very small employers are excluded from the law's requirements, and only in some states.²⁴⁵ Moreover, concern for cost or burden does not usually limit the workers' compensation claims of nonprofit employees or others who also may be doing socially valuable work.²⁴⁶ Thus, as was the case with regard to unrestricted subject payment, in order for cost to justify a refusal to protect injured subjects to the same extent as injured workers, it would have to be demonstrated that research progress is more critical than progress in other areas that might be affected by workers' compensation requirements. Bearing that possible caveat in mind, no convincing case exists thus far for differential treatment between the two groups.

2. Covered Harm and Causation

With these broad issues of scope, need, and burden resolved, the next question is whether there are any details of the workers' compensation system that would render a similar approach for injured subjects unworkable, since of course this would be a legitimate reason to protect subjects differently. Ultimately, some tweaking would be needed, but there is nothing so inherently different about the research context that a system of guaranteed and systematic compensation for injury ought to be rejected.

As for the types of harm that qualify for payment, workers' compensation seems to be a mixed fit for research. First, it fits well with regard to covering only those injuries that would result in documentable financial loss, and not

243 Serious research injuries are likely to be few and manageable, suggesting that the cost of care will be low and injuries will not keep subjects out of work for long. LEWIN GROUP, *supra* note 227, at ES-2, 32-33; Mariner, *supra* note 212, at 118; Pike, *supra* note 212, at 60-61; Resnick, *supra* note 31, at 265; Steinbrook, *supra* note 213, at 1873. Nonetheless, a system of guaranteed compensation for subjects would almost certainly be more expensive than the status quo, since no-fault compensation plans "may help reduce the number of large awards to subjects [awarded through the tort system] only at the expense of increasing the number of small awards." Resnick, *supra* note 31, at 284. But importantly, this equation does not account for how research is currently hindered by costly defensive behavior generated by fear of litigation that may never come to pass or be successful.

244 See Beh, *supra* note 225, at 12; LEWIN GROUP, *supra* note 227, at 1-2; *Moral Science*, *supra* note 17, at 67; Resnick, *supra* note 31, at 267.

245 *Workers Compensation Laws – State by State Comparison*, NAT'L FED'N OF INDEP. BUS., <http://www.nfib.com/legal-center/compliance-resource-center/compliance-resource-item/cmsid/57181> (last visited Nov. 22, 2013).

246 Bazil Facchina et al., *Privileges and Exemptions Enjoyed by Nonprofit Organizations*, 28 U.S.F. L. REV. 85, 117 (1993).

minor problems like transient nausea likely to occur in a great deal of trials.²⁴⁷ Second, coverage of harms without regard to fault appears to be appropriate for HSR, given that none of the reasons articulated above for cost shifting away from injured subjects depends on fault or the lack thereof.²⁴⁸ Moreover, no-fault systems can have a variety of benefits such as speed, cost-effectiveness, and congeniality that would be beneficial in the research setting, and nearly every country that sponsors, hosts, or conducts substantial amounts of research – other than the U.S. – has implemented a no-fault compensation system for research-related injury, demonstrating feasibility.²⁴⁹

However, the fact that workers' compensation covers harms arising out of and in the course of work is likely too broad, at least for some subjects. This is because those participating in therapeutic or prevention research might have experienced harms in clinical care that would have been similar in type or magnitude to those caused by research. Although these harms would meet the workers' compensation standard, compensating the subject looks like a windfall; this should not be a qualified harm.²⁵⁰ A related issue is that all effective therapies and preventative interventions have benefits and drawbacks, such that a subject in a therapeutic or prevention study might be benefited in one way and injured in another.²⁵¹ Ultimately, the workers' compensation standard for compensable injury is not equipped to handle this sort of problem because it is meant to remedy pure harms. Countervailing medical benefits and harms are likely to be common for HSR, however, and really cannot be ignored.

For subjects in therapeutic or prevention research, then, qualified harms must be more circumscribed than they are for purposes of worker's compensation. Rather than all injuries arising out of and in the course of research, qualified harms should be limited to those different or worse than what subjects could have expected in clinical care, and within that subset, limited to net harms.²⁵² In addition, net harms should not be discounted on the basis of financial benefits (*i.e.*, any payment a subject receives for participating), which is consistent with the fact that workers are compensated for work-related injury regardless of how much they make.

The research setting also poses other difficulties, but they are not necessarily unique. For example, when patients become subjects, it may be difficult to distinguish whether the symptoms experienced are compensable harms caused by the research or whether they are just the consequence of the subject's underlying

247 Pike, *supra* note 212, at 57.

248 Beh, *supra* note 225, at 12; *Moral Science*, *supra* note 17, at 63.

249 Pike, *supra* note 212, at 46.

250 See Adams & Shea-Stonum, *supra* note 228, at 642; Pike, *supra* note 212, at 46.

251 See *Moral Science*, *supra* note 17, at 69 (noting that it is necessary to determine whether compensable research injury should include side effects that follow an effective therapeutic intervention); VanderWalde & Kurzban, *supra* note 26, at 546 (questioning whether it is a harm when an experimental intervention lengthens life but causes other side effects).

252 Ackerman, *supra* note 28, at 3.

condition.²⁵³ However, a determination of research-relatedness must already be made for purposes of reporting adverse events to regulatory agencies and IRBs, so this problem is familiar and cannot be avoided.²⁵⁴ It is also already being addressed by the research institutions and handful of regulatory policies that currently provide or require compensation for study-related injury.²⁵⁵ Moreover, similar questions can arise for workers' compensation,²⁵⁶ particularly given its coverage of occupational disease that may be less obviously linked to the workplace because it develops over time and/or occurs in the general population as an ordinary disease of life.²⁵⁷ Recovering monetary costs in such cases is not necessarily easy,²⁵⁸ but the point is that factual issues of causation are not foreign to workers' compensation or to research. Thus, standing alone, difficulty in establishing causation for research-related injury is not a reason to treat injured subjects differently.²⁵⁹ And, importantly, for some research-related injuries, causation will be readily apparent.

3. Funding, Tort Preemption, and Additional Concerns

There remain a few additional reasons to suggest that even though injured subjects should be compensated, the workers' compensation mechanism might not be the right approach. First, perhaps the fact that medical research redounds to the public benefit²⁶⁰ suggests that unlike the workers' compensation system, which is funded by employers, the compensation system for injured subjects should be publicly funded.²⁶¹ But many companies perform work that is socially valuable and still remain directly responsible for workers' compensation coverage. There is also a very clear sense in which compensation for subjects' injuries will be paid by the public, despite being paid first by those conducting the research: if research is conducted with federal money, compensation would come out of tax revenues, and if research is privately-sponsored, the cost of compensation would be rolled into the prices of medical products.²⁶²

253 See Alexander M. Capron, *When Experiments Go Wrong: The U.S. Perspective*, 15 J. CLINICAL ETHICS 22, 25 (2004); LEWIN GROUP, *supra* note 227, at 33; IOM, *supra* note 213, at 193; Mariner, *supra* note 212, at 121; Pike, *supra* note 212, at 28, 56; Resnick, *supra* note 31, at 266; Steinbrook, *supra* note 213, at 1872; VanderWalde & Kurzban, *supra* note 26, at 546.

254 Pike, *supra* note 212, at 54-55.

255 See, e.g., *Moral Science*, *supra* note 17, at 122 n.117 (describing institutional compensation programs at the University of Washington and elsewhere).

256 See Resnick, *supra* note 31, at 266 (recognizing that courts and workers' compensation panels have to deal with complex causation problems on a routine basis).

257 Subjects' injuries may also be latent.

258 See generally STAFF OF H. COMM. ON WAYS & MEANS, *supra* note 221, at 2; Peirce & Dworkin, *supra* note 217, at 659.

259 See, e.g., Pike, *supra* note 212, at 56.

260 Mariner, *supra* note 212, at 121; *Moral Science*, *supra* note 17, at 58-60.

261 See Adams & Shea-Stonum, *supra* note 228, at 640-41.

262 See VanderWalde & Kurzban, *supra* note 26, at 545.

What about the issue of tort preemption? Some have suggested that injured subjects should have the option of accepting no-fault compensation and waiving their right to sue, or waiving their right to no-fault compensation and suing instead.²⁶³ This would give subjects who are likely to be undercompensated by the no-fault scheme a chance to do better, but it is unclear why subjects should get this opportunity when other workers do not. Perhaps the difference is that whereas subjects are unlikely to be at fault for their injuries, workers' compensation as the exclusive remedy for workplace injury despite employer fault was a concession to employers who would be responsible for payment despite employee fault.²⁶⁴ Nonetheless, the fact remains that injured subjects would be able to recover even in the absence of any fault whatsoever. Thus, those conducting research have a similar argument that they should not also have to face greater liability when they are in fact at fault.

In addition, there is a sense in which a single system for compensating injured subjects is fairer, since similarly injured subjects will be compensated similarly without regard to their sophistication to navigate the tort system, for example. Another potentially attractive feature of no-fault compensation paired with tort preemption is that it could increase subject trust and solidarity with researchers by sending a clear message: you will be taken care of, there is no need to fight to get what you are rightfully owed, and there is no room for adversity in the research relationship.²⁶⁵ There is some concern that an exclusive no-fault approach will be less able to deter bad behavior,²⁶⁶ but eliminating the specter of negligence liability may help minimize the defensive posture that has come to predominate the culture of HSR, reflected in everything from legalistic consent to overly nit-picky IRB review.²⁶⁷ That said, however, it would be appropriate to preserve the opportunity to sue in the event of intentional or egregious behavior, which some workers' compensation laws allow.²⁶⁸

263 See Beh, *supra* note 225, at 12, 13; Mariner, *supra* note 212, at 119; Pike, *supra* note 212, at 49; Resnick, *supra* note 31, at 283. This is how the National Vaccine Injury Compensation Program works. See *Vaccine Injury Compensation Program (VICP)*, NAT'L VACCINE INFO. CENTER, <http://www.nvic.org/injury-compensation.aspx> (last visited Nov. 22, 2013); Beh, *supra* note 225, at 13.

264 Adams & Shea-Stonum, *supra* note 228, at 639.

265 On the other hand, it is possible that tort preemption could have the unintended consequence of scaring away potential subjects who are wary of relinquishing their right to sue. And the workers' compensation system itself can become adversarial, as previously noted. See *supra* note 223 and accompanying text; see also ROBERT A. KAGAN, *ADVERSARIAL LEGALISM: THE AMERICAN WAY OF LAW* 130 (2001).

266 Adams & Shea-Stonum, *supra* note 228, at 643.

267 See IOM, *supra* note 213, at 189, 122, 178, 181; NBAC, *supra* note 225, at 13, 62, 117; see also Resnick, *supra* note 31, at 283 (explaining that a no-fault system may improve systems and procedures for preventing injuries by encouraging open communication).

268 See, e.g., *Workers' Compensation: Can I Sue My Employer Instead?*, FINDLAW, <http://injury.findlaw.com/workers-compensation/workers-compensation-basics-overview/workers-comp-sue-employer.html> (last visited Nov. 22, 2013); Workers' Compensation Agency, *Frequently Asked Questions*, MICH. DEP'T LICENSING AND REG. AFF., <http://www.michigan.gov/wca/0,4682,7->

Other possible reasons for differential approaches for injured workers and injured subjects can also be rejected. Most commentators maintain that subjects' consent to research participation cannot waive their moral right to compensation for injury,²⁶⁹ and although consent to work may now take the existence of workers' compensation into account as part of what is being agreed to, workers' consent did not previously vitiate their moral claims to implementation of such a system. It also does no work to point out that some subjects are motivated by their own self-interest,²⁷⁰ since the same is true of workers. Finally, there is no more need for concern that a compensation system would be viewed as a license to embark on riskier research than it would be viewed as a license to permit more dangerous workplaces,²⁷¹ and the same safety and review standards would remain in place regardless of compensation for injury.

4. Unpaid Subjects

Although the preceding discussion has laid the case for compensating injured subjects and injured workers similarly, one important question remains. Since unpaid workers – volunteers – are often excluded from workers' compensation coverage as nonemployees, should the same be true for unpaid subjects?

There are generally two types of unpaid subjects, the pure altruist and the subject induced to participate by nonmonetary benefits such as possible therapeutic improvement. The latter resembles the unpaid intern seeking professional advancement in the sense that both are volunteering primarily for their own purposes and are remunerated in an important but non-monetary way. Moreover, both may make important contributions to the projects in which they are involved.²⁷² Some states specifically exclude interns from workers' compensation coverage, but elsewhere, courts and administrative boards have granted them benefits, which seems to be the right approach given interns' likely lack of bargaining power and the problems with tort litigation described above.²⁷³ That same protection should apply to the unpaid benefit-seeking subject,

191-27210-41833--F,00.html (last visited Dec. 4, 2013); *Workplace Injury: When You Can Sue Outside of Workers' Compensation*, NOLO, <http://www.nolo.com/legal-encyclopedia/workplace-injury-lawsuit-sue-30334.html> (last visited Nov. 22, 2013).

269 Adams & Shea-Stonum, *supra* note 228, at 609; James F. Childress, *Compensating Injured Research Subjects: The Moral Argument*, 6 HASTINGS CENTER REP. 21 (1976); IOM, *supra* note 213, at 191; *Moral Science*, *supra* note 17, at 58-60; Pike, *supra* note 212, at 20-21; Scott, *supra* note 214, at 421. *But see* Steinbrook, *supra* note 213, at 1872 (noting the contrary view that "routine compensation is not required because subjects are made aware of the risks through the informed-consent process, understand them, and voluntarily agree to participate").

270 See Beh, *supra* note 225, at 12.

271 See Adams & Shea-Stonum, *supra* note 228, at 641 n.95-96.

272 But note that this may necessitate payment under the Fair Labor Standards Act. See *supra* note 167 and accompanying text.

273 Mark Schappert, *Employers' Internship Toolkit* (2005), LE MOYNE C. (Aug. 2005), http://www.lemoyne.edu/Portals/11/pdf_content/career_services/INTERNSHIP_TOOLKIT.pdf.

considering that all of the same rationales described above for offering compensation for research-related injury hold true, and also because these subjects likely depend on the research in some way that renders it difficult for them to negotiate for such compensation on their own.²⁷⁴

On the other hand, altruistic unpaid subjects pose an interesting paradox. They could presumably walk away from research opportunities with ease if they (1) understood the likelihood and potential severity of research injury, and that they would be on their own in the event injury occurs, and (2) found that to be unacceptable. This suggests that no intervention is really needed, although some have questioned whether the first assumption is really true.²⁷⁵ But even if it is, failing to mandate compensation when these subjects are injured seems to contradict the fact that altruists have the greatest moral claim to such compensation – after all, they are taking on risks and burdens *exclusively* for others. Shouldering the cost of injury may be part of the gift such altruists are willing to make, but it does not seem reasonable to expect altruists to either make such a sacrifice or refrain from research participation, especially when others who made lesser sacrifices (in the sense of having greater self-interest) would be provided compensation. Then again, this is precisely what happens in some cases. For example, volunteers helping to build a house for Habitat for Humanity might have to sign a waiver and release of liability explicitly recognizing that they are not covered by workers' compensation insurance, even though a construction employee doing exactly the same thing while getting paid would be covered.²⁷⁶

What to do, then, with the true volunteer subject? Although it would not be outrageous or inconsistent to exclude them from mandatory no-fault compensation for injury,²⁷⁷ the more attractive option is to treat them like other volunteers who perform risky work that is highly socially valuable, offering protection as a matter of justice rather than to remedy inadequate bargaining power. For example, some states have chosen to recognize the "unselfish service" of volunteer emergency workers by covering them under workers' compensation laws.²⁷⁸ Given that altruistic volunteer subjects would face the

274 Mariner, *supra* note 212, at 123.

275 See Pike, *supra* note 212, at 44-45 (arguing that research subjects are not aware of the extent to which they are legally unprotected in the event of injury); see also Mariner, *supra* note 212, at 117.

276 See, e.g., *Waiver of Liability*, HABITAT FOR HUMANITY SMITH COUNTY, <http://www.smithcountyhabitat.org/volunteer-waiver-of-liability.html> (last visited Nov. 22, 2013). However, some organizations will voluntarily offer some protection to volunteer workers. See, e.g., *Frequently Asked Questions, Volunteering: General Questions*, HABITAT FOR HUMANITY GREATER BOS. <http://www.habitatboston.org/faq.html> (last visited Nov. 22, 2013).

277 James G. Hodge, Jr. et al., *The Legal Framework for Meeting Surge Capacity Through the Use of Volunteer Health Professionals During Public Health Emergencies and Other Disasters*, 22 J. CONTEMP. HEALTH L. POL'Y 5, 51 (2005).

278 See, e.g., *Volunteer Firefighter's and Volunteer Ambulance Worker's Guide to Workers' Compensation Benefits*, N.Y. WORKERS' COMP. BD., <http://www.wcb.ny.gov/content/main/>

same difficulties recovering in tort as other injured subjects, and given the potential value of limiting liability in the research setting, as well as the symbolic importance of protecting those engaged in socially valuable activities, all types of research subjects – paid and unpaid, healthy and patient, enrolled in therapeutic and nontherapeutic research – should be entitled to at least the same benefits guaranteed to injured workers under the workers’ compensation system: no-fault compensation for at least the costs of medical care and lost wages resulting from their injuries. The HSR regulations should be revised accordingly.²⁷⁹

C. Working Conditions and Inspections

With substantial changes in order for subject compensation, next consider the differential extent to which the HSR regulations and those governing the workplace address day-to-day working conditions. At first glance, the difference seems striking, with employees granted a variety of more specific protections. But upon closer inspection, little regulatory change is needed for human subjects.

For obvious reasons, working conditions are of greatest concern to those subjects enrolling in confinement studies at inpatient research centers. These sites differ substantially with regard to quality of facilities, amenities, and staff,²⁸⁰ with the worst facing problems described starkly by one self-titled “professional guinea pig”: “[o]vercrowding, no hot showers, sleeping in an easy chair, incredibly cheap shit for dinner, creepy guys from New York jails—all these are a poor man’s worries Where are these things in the regulators’ paperwork?”²⁸¹

Although IRBs and HSR regulatory agencies have the authority to inspect trial sites, such inspections are rare (largely due to resource constraints). When they do occur, they tend to focus on things like recordkeeping, protocol deviations, and informed consent issues.²⁸² Paper-based review of research is the

firefighters/VFAWLawIntro.jsp (last visited Nov. 22, 2013); see also James G. Hodge, Jr. et al., *A Hidden Epidemic: Assessing the Legal Environment Underlying Mental and Behavioral Health Conditions in Emergencies*, 4 ST. LOUIS U. J. HEALTH L. & POL’Y 33, 66 (2010); Rubinstein, *supra* note 86, at 182 n.174.

279 In addition to a new provision in the HSR regulations mandating compensation for injury, the prohibition on exculpatory language would also need to be modified in light of tort preemption. Other legal changes may also be necessary as a result of the Anti-Deficiency Act and the Adequacy of Appropriations Act. For a discussion of these issues, see Pike, *supra* note 212, at 47-49, 59.

280 See Elliott, *supra* note 30; GUINEA PIG ZERO, *supra* note 27, at 16-28; JIM HOGSHIRE, SELL YOURSELF TO SCIENCE: THE COMPLETE GUIDE TO SELLING YOUR ORGANS, BODY FLUIDS, BODILY FUNCTIONS AND BEING A HUMAN GUINEA PIG 9, 20-25 (1992); Ondrusek, *supra* note 27, at 9-10, 113-119.

281 Elliott, *supra* note 30 (quoting Robert Helms).

282 See FDA et al., *Guidance for Industry: Oversight of Clinical Investigations – A Risk-Based Approach to Monitoring*, HHS 3 (Aug. 2013), <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM269919.pdf> (“Monitoring activities include communication with the . . . study site staff; review of the study site’s processes, procedures, and records; and verification of the accuracy of data submitted to the sponsor.”);

norm, but most protocols contain little information on the physical aspects of the sites where research is conducted, and IRB members may not be independently familiar with them.²⁸³ Investigators are charged by FDA with generally protecting the rights, safety, and welfare of research subjects,²⁸⁴ but they have substantial discretion with regard to site conditions, which are not explicitly addressed by the HSR regulations. The same is true for study monitors selected by sponsors to oversee the conduct and progress of clinical investigations at the site level.²⁸⁵

In contrast, although the Occupational Safety and Health Administration (OSHA) also inspects a very small percentage of workplaces each year,²⁸⁶ its regulations and overall focus are heavily concerned with specific conditions on the ground.²⁸⁷ In fact, OSHA has a variety of requirements for the living quarters at temporary labor camps,²⁸⁸ as well as housing for certain agricultural workers.²⁸⁹ Moreover, its general duty clause requires employers to provide a place of employment “free from recognizable hazards that are causing or are likely to cause death or serious harm to employees.” This requirement would be violated by an employer who does nothing to prevent or abate a recognized hazard of workplace violence, such as threats, intimidation, or other indicators.²⁹⁰

Of course, trial sites are not as risky in and of themselves as mines, factories, or other dangerous worksites, so similarly stringent inspection requirements would seem overly burdensome. But is the difference between the OSHA and HSR standards and regulations themselves problematic? Not really. First, to the extent that sites are the workplaces of uncontroverted employees – research staff – basic OSHA standards regarding things like means of egress, ventilation, sanitation, and fire protection²⁹¹ must already be satisfied, and research subjects

Bioresearch Monitoring (BIMO) Metrics – FY '10, FDA (2010), <http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/RunningClinicalTrials/UCM261855.pdf>; Daniel R. Levinson, *The Food and Drug Administration's Oversight of Clinical Trials*, HHS (Sept. 2007), <http://www.oig.hhs.gov/oei/reports/oei-01-06-00160.pdf>; OHRP, *OHRP's Compliance Oversight Procedures for Evaluating Institutions*, HHS (Oct. 14, 2009), <http://www.hhs.gov/ohrp/compliance/evaluation/ohrpcomp.pdf>; ABADIE, *supra* note 27, at 155.

283 Elliott, *supra* note 30; Ondrusek, *supra* note 27, at 48-49, 113-119.

284 21 C.F.R. § 312.60 (2013).

285 21 C.F.R. § 312.53(d). Note that the Common Rule does not impose specific obligations on sponsors or investigators, unlike the FDA regulations.

286 See *Indicator 18: Occupational Safety and Health Administration (OSHA) Enforcement Activities*, MD. DEP'T. HEALTH & MENTAL HYGIENE, <http://phpa.dhmd.maryland.gov/OEHFP/EH/SitePages/Health-Indicator-18.aspx> (last visited Nov. 22, 2013) (finding that OSHA inspected only 1.2% of all establishments under its jurisdiction in 2000).

287 See generally 29 C.F.R. pt. 1910 (2013).

288 29 C.F.R. § 1910.142.

289 29 C.F.R. § 500.130.

290 Occupational Safety & Health Admin. (OSHA), *Workplace Violence: Enforcement*, U.S. DEP'T LAB., <http://www.osha.gov/SLTC/workplaceviolence/standards.html> (last visited Nov. 22, 2013).

291 See generally 29 C.F.R. § 1910.

would be incidental beneficiaries of compliance and enforcement. Second, the health care facilities that may be used for research are heavily regulated,²⁹² and even if research is conducted at a stand-alone site not otherwise subject to specialized health care facility regulations, the site will still be subject to local building safety, sanitation, and fire codes. Indeed, it was violation of these requirements that led to the 2005 demolition of a trial site in Miami.²⁹³ Thus, the HSR regulations do not necessarily need to impose their own site-specific standards, which may be redundant, nor should HSR regulators step in to enforce standards that really ought to be enforced by other agencies.

Instead, the best approach to ensuring commensurate protection for workers and human subjects is to make sure that the designated authorities appropriately enforce the *existing* health and safety requirements applicable to research sites. It may also be necessary to ensure that stand-alone research sites are covered by appropriate health care facilities regulations, if they are not already. And it would be a reasonable compromise for the HSR regulations to require that IRBs receive certification or explicit assurance that facilities used for research purposes are in compliance with all applicable facilities codes, as OSHA does in some contexts.²⁹⁴ Finally, at an absolute minimum, the HSR regulations should make clear that the responsibility for protecting subject welfare includes not only consideration of study interventions but also study conditions, and should also explicitly protect subjects against retaliatory action for reporting violations.²⁹⁵ Some site level issues will certainly remain, since there is no worklaw requirement that supervisors be nice, that the food be good, or that entertainment options be provided (all of which are complaints that have been lodged by the subject community). However, these can be appropriately left to the market as they are far removed from actual safety concerns, and instead reflect the sorts of discomforts that are relatively commonplace in daily life.

D. Collective Bargaining and Unions

As a final point of comparison, consider that employees also have greater protection than research subjects with regard to their rights to concerted action and collective bargaining under the National Labor Relations Act (NLRA). At

292 See, e.g., FLA. ADMIN. CODE ANN. r. 59A-3.081 (2013) (Physical Plant Requirements for General, Rehabilitation and Psychiatric Hospitals); 25 TEX. ADMIN. CODE § 1.133, Subch. 1 (2013) (Hospital Licensing Physical Plant and Construction Requirements); JOINT COMMISSION, 2012 HOSPITAL ACCREDITATION STANDARDS (2011).

293 David Evans, Michael Smith, & Liz Willen, *Big Pharma's Shameful Secret*, BLOOMBERG MARKETS, Dec. 2005 (describing study conditions at SFBC International's Miami site); David Evans, *SFBC Ordered to Demolish Miami Drug Test Center*, BLOOMBERG (May 18, 2006, 4:10 PM), <http://www.bloomberg.com/apps/news?pid=newsarchive&sid=a3FrIL6PPXkc> (describing conditions resulting in the site's demolition).

294 See, e.g., 29 C.F.R. § 500.135 (2013).

295 See, e.g., *Safety and Health Standards: Occupational Safety and Health*, U.S. DEP'T LAB., <http://www.dol.gov/compliance/guide/osh.htm> (last updated Sept. 2009).

first glance, it may appear that subjects are just as well protected via IRBs, if not more so, given that IRBs stand as the constant intermediary protecting the interests of subjects against those of researchers and sponsors. Further, although researchers and sponsors have no choice but to engage with IRBs, IRBs are under no obligation to “bargain” with those conducting the research; unlike bargaining representatives under labor law, IRBs can command and compel because their approval is legally required before research can proceed or continue.

It would be a mistake, however, to view IRBs as actually representing subjects in the way that unions represent employees. With the exception of research involving prisoners, the regulations do not require that any subjects be included on the board, or that subjects be consulted at all, and no one is charged with speaking on the subjects’ behalf.²⁹⁶ The goal is only to protect them, and a mission to protect can be quite different from a mission to improve. Moreover, even though a subject can always bring his or her concerns to the IRB, there is no guarantee that the IRB’s response will be what a subject had in mind; a complaint about low payment is unlikely to result in higher rates if a board is concerned about undue inducement, for example.

On the other hand, there is a growing trend toward community engagement in HSR, with the goal of providing communities greater ownership of and information about research projects.²⁹⁷ Thus, a number of trial networks and research sites now work with Community Advisory Boards (CABs) as a way to provide those affected by or involved in research at the local level a way of voicing their needs and concerns.²⁹⁸ Even this, however, is a far cry from the protection offered to employees by the right to bargain collectively. First, the community in question may include anyone bearing a stake in the research, from government representatives to patient advocacy groups to health care workers to subjects and their families.²⁹⁹ Second, and more importantly, while it may be good ethical practice to engage in community consultation, the regulations do not

296 45 C.F.R. § 46.107; 21 C.F.R. § 56.107 (describing IRB membership requirements). For research involving prisoners, at least one member of the Board must be a prisoner or prisoner representative. 45 C.F.R. § 46.305.

297 See, e.g., 1 NAT’L BIOETHICS ADVISORY COMM’N, *ETHICAL AND POLICY ISSUES IN INTERNATIONAL RESEARCH: CLINICAL TRIALS IN DEVELOPING COUNTRIES*, 30 (2001); UNAIDS, *Good Participatory Practice: Guidelines for Biomedical HIV Prevention Trials* (June 2011), http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC1853_GPP_Guidelines_2011_en.pdf.

298 See, e.g., Cmty. Recommendations Working Grp., *Recommendations for Community Involvement in National Institute of Allergy and Infectious Diseases HIV/AIDS Clinical Trials Research*, HIV VACCINE TRIALS NETWORK (Feb. 2009), http://www.hvtm.org/community/CAB_Recommendations_Certified.pdf.

299 *Id.* at 6.

require it outside the context of emergency research conducted without subject consent.³⁰⁰

Ultimately, there is no requirement that anyone actually hear and consider the requests and demands of research subjects. If a subject attempted to negotiate greater benefits for those enrolled in a trial, he would not be entitled to legal protection against retaliation.³⁰¹ Even if he were not penalized, researchers and sponsors would be under no obligation to negotiate with him or any other representative, and even if they wanted to, any agreements reached could be thwarted by the IRB. Then again, subjects would be under no obligation to follow any standards whatsoever in their negotiations.

Employees covered by the NLRA, in contrast, have the “right to self-organization, to form, join, or assist labor organizations, to bargain collectively through representatives of their own choosing, and to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection...”³⁰² Employers are prohibited from interfering with, restraining, or coercing employees in the exercise of these rights.³⁰³ The Act also protects both employees and employers against certain behaviors by labor organizations.³⁰⁴

To engage in collective bargaining, employees must come together to determine that they want a bargaining representative, and select one. If and once they have done so, both the employer and the representative have a legal obligation to bargain in good faith about wages, hours, and other conditions of employment,³⁰⁵ although this does not compel either party to agree to any proposal made by the other.³⁰⁶ Moreover, employees may strike (and employers may impose a lockout) in order to further their position, so long as they (or their representatives) remain engaged in good faith bargaining throughout.³⁰⁷

Employees are also protected in their right to *refrain* from organizing, joining a union, bargaining collectively, or engaging in other concerted

300 21 C.F.R. § 50.24(a)(7)(i); FDA et al., *Guidance for Institutional Review Boards, Clinical Investigators, and Sponsors: Exception from Informed Consent Requirements for Emergency Research*, HHS 25-34 (Apr. 2013), <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM249673.pdf>.

301 As one website offering tips for subjects advises: “Don’t circulate petitions (or sign them) protesting your pay, or restrictions on the unit . . . Usually this behaviour will get you banned from the clinic in the future.” *Tips for Clinical Trials and Clinical Study Volunteers*, GPGP.NET (2009), <http://www.gpgp.net/tips>.

302 29 U.S.C. § 157 (2006).

303 29 U.S.C. § 158(a)(1).

304 29 U.S.C. § 158(b).

305 29 U.S.C. §§ 158(a)(5), (b)(3), (d); *Basic Guide to the National Labor Relations Act: General Principles of Law Under the Statute and Procedures of the National Labor Relations Board*, NAT’L LABOR RELATIONS BD. (NLRB) 20 (1997), <https://www.nlrb.gov/sites/default/files/documents/224/basicguide.pdf>.

306 29 U.S.C. § 158(d).

307 *Basic Guide to the National Labor Relations Act*, *supra* note 305, at 3-5, 16.

activities.³⁰⁸ However, because a bargaining representative need only be designated by the majority of the employees in a bargaining unit, an individual who would have selected a different representative or chosen not to engage in collective bargaining at all is nonetheless stuck; all employees will be bound by any collective bargaining agreement negotiated by the representative.³⁰⁹ Another restriction on the “right to refrain” is that although applicants for employment cannot be required to be members of a union in order to be hired, and employees cannot be required to join or maintain membership in a union in order to retain their jobs, the NLRA does permit “union security” agreements in some cases. Thus, a union and an employer can make an agreement that requires employees to pay their share of “financial core” costs relating to the union’s representational activities (short of actual membership) in order to retain their jobs.³¹⁰ However, in a “right-to-work” state, which now describes nearly half of the U.S., employees may not be forced to even financially support a union.³¹¹

All things considered, even though subjects may have some greater freedoms than employees when it comes to negotiating, these are outweighed by the ability of employees to come together to achieve things they likely could not on their own, and to have a voice that *must* be heard by those in charge.³¹² But should subjects be leveled up with regard to collective bargaining and unionization, as some bioethicists have argued?

The idea appears intuitively appealing, considering the similarities between subjects and workers and the challenges they might face. For example, although the stated function of the NLRA is to help mitigate industrial strife that would otherwise burden or obstruct commerce if left unchecked, it also recognizes the “inequality of bargaining power between employees who do not possess full freedom of association or actual liberty of contract” and seeks to “restor[e] equality of bargaining power between employers and employees” as a way to protect commerce.³¹³ The fact that research subjects are not currently guaranteed NLRA-type rights has clearly not brought commerce – or research – to a halt, so it may seem unnecessary to institute any change in this regard. However, that may be a testament to just how poor subjects’ bargaining power truly is, rather

308 29 U.S.C. § 157.

309 29 U.S.C. § 159(a); *Basic Guide to the National Labor Relations Act*, *supra* note 305, at 8.

310 *Basic Guide to the National Labor Relations Act*, *supra* note 305, at 2.

311 *Right to Work States*, NAT’L RIGHT TO WORK LEGAL DEF. FOUND., <http://www.nrtw.org/rtps.htm> (last visited Nov. 6, 2013); “*Right to Work*” for Less, AFL-CIO, <http://www.aflcio.org/Legislation-and-Politics/State-Legislative-Battles/Ongoing-State-Legislative-Attacks/Right-to-Work-for-Less> (last visited Nov. 22, 2013).

312 This is not to suggest that the NLRA rights are perfect or perfectly enforced, and in fact many labor scholars have pointed out deficiencies and proposed solutions. *See, e.g.*, Stephen F. Befort, *Labor and Employment Law at the Millennium: A Historical Review and Critical Assessment*, 43 B.C. L. REV. 351, 433-443 (2002). However, the rights themselves are more than what subjects have at present.

313 29 U.S.C. § 151 (2006).

than an indication that they are satisfied with the terms and conditions of the research opportunities they are offered. With the exception of true altruists, individual subjects generally hold the weakest position in the HSR relationship, and their position could be improved by collective action on a large enough scale. Thus, although the underlying purpose of the NLRA does not necessarily dictate extension of similar rights to research subjects, it certainly does not conflict with such extension.

Consider, however, the obstacles that would stand in the way of collective bargaining and unionization by most research subjects. With some limited exceptions,³¹⁴ studies are usually developed and approved before any interaction with subjects occurs, and before subjects are even identified, so there is not yet anyone for those conducting the research to come together and bargain with, let alone anyone whose interests can be represented, except in the abstract. Subjects could decide to organize for improved terms and conditions once they have been enrolled, at which point they have improved bargaining power together given the sunken investment in their data, but by then, it is quite late to make any changes. More importantly, outside of confinement studies, they may never be in the same place at the same time or even know who else is in a study, which obviously makes the identification of common interests and the push to organize difficult. In addition, subjects may participate in research only once or just a few times, such that even though they have shared interests, they have no long-term investment in the hassle of organizing.³¹⁵ Even if they could be assured protection against retaliation for concerted action, in order for such action to impose the desired bargaining pressure, the subjects would have to be willing to take some risk together, *e.g.*, to strike until their demands are met. Subjects who just want the money or especially the potential therapeutic benefits may be unwilling to do so.³¹⁶ Moreover, bargaining study-by-temporary-study would be terribly inefficient. If organization and bargaining had to occur at this level, it is very unlikely that the burdens would outweigh the benefits, and that enough subjects would have adequate motivation to take the necessary steps.

Given their greater durability than single studies, the research site or sponsor would seem to be the preferable locus of organizational activity, but that assumes there are some subjects who have a site- or sponsor-level interest, beyond single studies. That may be true for a subset of repeat and professional subjects, although their numbers may be inadequate to provide significant leverage, particularly given the difficulties they are likely to face identifying one another

314 For example, consider *Greenberg v. Miami Children's Hosp. Research Inst.*, 264 F. Supp. 2d 1064 (S.D. Fla. 2003), where patients themselves initiated research on their rare disease and would likely have had relatively strong bargaining power (had they attempted to reach a clear agreement *ex ante*) by virtue of having something rare that researchers were interested in.

315 See Befort, *supra* note 94, at 170 (noting that many contingent workers do not see the benefits of union representation in the context of short-term employment).

316 ABADIE, *supra* note 27, at 58; *see also* GALL, *supra* note 75, at 110-111 (explaining the similar problem of organizing some commercial sex workers).

and/or staying in contact once any given study is over.³¹⁷ And even if they successfully negotiate terms to cover all studies at a given site, for example, there is the problem of free-riding, unless all subjects enrolling at that site are required to contribute in some way to the negotiating group.

Subjects' best bet would be to "organize the industry" along the lines of the Screen Actors Guild (SAG) and other performing arts unions,³¹⁸ but this too would be tremendously challenging. Like subjects, actors are engaged in short-term projects with a variety of different employers. To deal with these unique circumstances, SAG negotiates basic contracts with producers to establish minimum wages and working conditions for actors.³¹⁹ Producers who want to hire SAG members must agree to the terms of a SAG contract, and actors who want to perform in a "Guild Signatory production" must generally either be SAG members or join within a certain period of time.³²⁰ Moreover, SAG members are bound by "Global Rule One," which dictates that "no member shall work as a performer or make an agreement to work as a performer for any producer who has not executed a basic minimum agreement with the Guild which is in full force and effect."³²¹ Failure to abide by this rule can lead to disciplinary action including expulsion.

Ultimately, actors want to be SAG members because otherwise their job prospects are limited and producers want to be Guild Signatories because otherwise their talent prospects are limited. Note that for the system to work, there must be a critical mass on at least one side of the equation: enough SAG members refusing to work for non-SAG producers that the producers are strongly motivated to deal, or enough SAG producers refusing to hire non-SAG members (within the limits of union security agreements) that the actors are strongly motivated to join. On the other hand, if enough producers refused to become Guild Signatories, the actors could work without joining SAG – in fact, SAG membership would hold them back because of Global Rule One – and if enough

317 Helms, *supra* note 5 ("The guinea pig workforce may be too fragmented and fluid to form even an unofficial union."); ABADIE, *supra* note 27, at 82-83 (explaining that once a trial is over, subjects usually do not remain in contact).

318 Others have also relied on the SAG model to demonstrate how diverse and nontraditional workers who might be difficult to unionize could successfully do so. See, e.g., Patricia Ball, Comment, *The New Traditional Employment Relationship: An Examination of Proposed Legal and Structural Reforms for Contingent Workers from the Perspectives of Involuntary Impermanent Workers and Those Who Employ Them*, 43 SANTA CLARA L. REV. 901, 938 (2002).

319 *Acting in Your Interest*, SCREEN ACTORS GUILD (SAG) 15 (2011), http://www.sag.org/files/sag/documents/Acting_In_Your_Interest_0.pdf.

320 *Getting Started as an Actor FAQ: What Is a "Guild Signatory?"*, SAG-AFTRA, <http://www.sagaftra.org/content/getting-started-actor-faq> (last visited Nov. 22, 2013). It is important to recognize that this is a union security provision, however, and therefore is not absolute. See Bob Labate, *The Mystery of Financial Core*, HOLLAND & KNIGHT (Sept. 15, 2000), <http://www.hklaw.com/publications/The-Mystery-of-Financial-Core-09-15-2000/>; *Get the Facts About Financial Core*, SAG-AFTRA <http://www.sagaftra.org/get-facts-about-financial-core> (last visited Nov. 22, 2013).

321 *Acting in Your Interest*, *supra* note 319, at 5.

actors refrained from membership, producers would have no reason to become Guild Signatories.

Putting this all together, in order for research subjects to engage in successful collective action to negotiate better terms for study participation, they could wait until a study is underway and then jointly threaten to quit – an unlikely scenario. Or they could adopt the SAG model by forming a union and requiring members to refuse to enroll in any study not covered by a union contract, which would include union security provisions. However, if sponsors and sites are able to enroll enough non-union subjects, they have no reason to sign a union contract, and the whole thing falls apart. Thus, the essential element is encouraging the critical mass of subjects to join. But is that possible?

Assuming that the right to refrain from union membership would be retained in the realm of HSR, which undercuts unions but preserves important individual freedoms, the most that could be required of subjects seeking enrollment in union studies would be satisfaction of financial core obligations to the union, not actual membership.³²² In that case, fee-paying nonmember subjects could participate in union studies, as well as non-union studies because they would not be bound by the requirement imposed on members to refuse enrollment in non-union work. Moreover, in “right-to-work” states, even financial core status cannot be required.³²³ Since this clearly allows free-riding on the union’s efforts,³²⁴ there are good arguments that it should not be allowed for HSR (or other types of work). However, even though there is not necessarily any legal right to be included in a research study, there might be a strong moral argument for treating participation in at least therapeutic studies as a “right-to-work” endeavor, the opportunity for which should not be denied on the basis of non-membership in a subjects’ union, or non-payment, as the case may be.³²⁵ Indeed, research participation may be a patient’s best hope for care, particularly if they have exhausted all other options and face a serious disease. Similarly, union membership itself would be problematic if it would technically preclude participation in a desirable, non-union therapeutic study through something like Rule One. Subjects likely to enroll in only a single study or for whom the therapeutic stakes are very high – namely, those for whom the worker analogy is the weakest – would obviously choose enrollment over union membership in this

322 See *Comm’n Workers v. Beck*, 487 U.S. 735, 745 (1988); *N.L.R.B. v. Gen. Motors Corp.*, 373 U.S. 734, 742 (1963).

323 See Michael M. Oswalt, Note, *The Grand Bargain: Revitalizing Labor through NLRA Reform and Radical Workplace Relations*, 57 DUKE L.J. 691, 700 (2007).

324 *Id.* at 701; see also Susan Guyett, *Indiana Becomes 23rd “Right-to-Work” State*, REUTERS (Feb. 1, 2012, 5:29 PM), <http://www.reuters.com/article/2012/02/01/us-unions-indiana-righttowork-idUSTRE81018920120201>.

325 See, e.g., Martha L. Elks, *The Right to Participate in Research Studies*, 122 J. LABORATORY & CLINICAL MED. 130 (1993) (arguing that it is important to protect subjects not only from the risks of research participation, but also to protect their interest in the benefits of research participation).

case. And subjects motivated purely by altruism – also in the weak analogy camp – would likely see no need to come together for mutual protection in this way. Ultimately, each of these factors would conspire to weaken any attempt at subject unionism.

Repeat and professional subjects are the most likely candidates for successful unionization, as they have the most “skin in the game” and are the most likely to benefit from the entertainment industry approach. Given their concentration in nontherapeutic, early phase studies, there seems to be no moral argument in favor of a right-to-work approach, although as a scientific matter, limiting participation in any type of research to union members (or those paying membership fees) could be problematic in terms of generalizability and bias. Thus, it seems that union security agreements would be unacceptable across the board in the research context.

There are also a number of other barriers to achieving the critical mass of subjects necessary to get any sites or sponsors to agree to collective bargaining contracts. As noted above, subjects may be unwilling to refrain from enrollment or participation as an exercise of power intended to push a deal through, especially because there may be enough casual or one-time subjects available to replace them. Moreover, if HSR regulations are amended to lift payment restrictions and guarantee subjects a minimum wage, compensation for injury, and decent working conditions, some of their chief complaints will have been addressed, potentially cutting down on their motivation to organize.³²⁶ And although the Internet could help,³²⁷ there is the fundamental problem of organizing a group of relatively transient, potentially stigmatized subjects who may fear running in opposition to a significant source of their income.³²⁸ Perhaps most importantly, if critical steps are taken to preserve the integrity, generalizability, and validity of research results by limiting repeat participation,³²⁹ there will be even fewer subjects who would care enough to organize.

So what is the bottom line for the HSR regulations – are the various obstacles to collective action sufficient reason not to level subjects up? At the very least, subjects should have the freedom to face these obstacles head on, and therefore ought to have a protected right to engage in concerted activity, if they so choose, just like workers protected by the NLRA.³³⁰ In the labor context, this right is

326 See ABADIE, *supra* note 27, at 58 (noting that subjects’ pay is generally good compared to other opportunities, so subjects have reduced incentive to challenge industry).

327 Richard B. Freeman & M. Marit Rehavi, *Helping Workers Online and Offline: Innovations in Union and Worker Organization Using the Internet*, (Nat’l Bureau of Econ. Research, Working Paper No. 13850, 2008), <http://www.nber.org/papers/w13850>.

328 See GALL, *supra* note 75, at 92-93, 189-218 (describing a number of barriers to organizing sex workers that apply with equal force to HSR).

329 See notes 187-190, *supra*, and accompanying text.

330 See Eli Naduris-Weissman, *The Worker Center Movement and Traditional Labor Law: A Contextual Analysis*, 30 BERKELEY J. EMP. & LAB. L. 232, 257 (2009).

quite broad,³³¹ as it should be for HSR. For example, sites and sponsors should be prohibited from discriminating against a subject or potential subject in enrollment or continued participation on the basis of that subject's present or former attempts to advocate on behalf of subjects, draw attention to subjects' concerns, inform subjects of their common interests and what they might be able to accomplish working together, and organize subjects for collective activity. Without these basic protections, subjects' bargaining power is limited to a degree beyond their inherent financial and/or therapeutic vulnerability. Of course, these concerted activities may fail to achieve subjects' goals, especially because individual subjects should also be free to avoid joining in, and those conducting the research would not be legally forced to give in to subjects' demands. But at least by protecting the right to act together or call others to action, subjects would have a good chance of being heard. Subjects aggrieved in their attempts to exercise this right should have recourse to the IRB and/or federal regulatory agencies overseeing HSR.³³²

The question of whether subjects ought to also be granted the right to bargain collectively through representatives of their own choosing is somewhat more complicated, however. Unlike protecting concerted activity, which would involve only the cost of enforcement, leveling up as to representative collective bargaining would involve the administrative costs and practical considerations of identifying appropriate bargaining units and conducting elections to determine whether a majority of subjects wish to unionize, and if so, who they want to represent them.³³³ Then again, these costs would only have to be incurred when the subjects have some chance of success, as when they can demonstrate that some threshold number supports representation, in which case the costs may be justified.³³⁴ The alternative path of persuading a site or sponsor to voluntarily recognize a subject bargaining representative after a showing of majority support seems preferable,³³⁵ but of course, may be difficult to achieve without the intervention of some third party.

Beyond the issue of resources, there may also be some concern that a collective bargaining approach would create a situation in which those conducting the research felt less responsibility for subject welfare, either because subjects would be viewed as adversaries³³⁶ or because subjects have a better means of protecting themselves. On the other hand, those conducting the research

331 *Id.* at 259-61.

332 Similarly, the NLRA offers no private right of action.

333 *What We Do: Conduct Elections*, NLRB <https://www.nlrb.gov/what-we-do/conduct-elections> (last visited Nov. 22, 2013).

334 For example, before the NLRB will hold a representation election, at least 30% of employees in the relevant bargaining unit must have signed a petition showing interest. *Id.*

335 *Id.*

336 GUINEA PIG ZERO, *supra* note 27, at 6 ("If we lab rats were to be taken under the regular labor laws and tried using traditional organizing methods, we'd end up with fewer freedoms, making less money, and we'd be lied to by the scientists more often.").

would have to abide by the terms of any collective bargaining agreement, and at a more fundamental level, the very reason collective bargaining would have been necessary is that subjects felt others were not adequately protecting their interests.³³⁷

Ultimately, failure to grant subjects a right to representative collective bargaining may not be a substantial problem in practice, since they are likely to face so many obstacles to organizing in the first place, including the need to adopt a right-to-work approach in the research setting for the reasons described above. However, if they could successfully organize, this protection would be essential to getting those who conduct research to engage in good faith bargaining as a matter of course, rather than forcing subjects to threaten a strike in order to even get sites and sponsors to the bargaining table. Many other types of workers also face difficulty exercising their labor rights, and labor unions are on the decline in general,³³⁸ but this is not a reason to eliminate those rights or not to offer them in the first place. If anything, it is a reason to seek ways to make them more accessible. Further debate is in order, but for now, there appears to be no compelling reason to refrain from leveling up by granting subjects the full gamut of NLRA rights.³³⁹ The best mechanism for doing so remains an open question,³⁴⁰ but the first step of protecting subjects' concerted activity should be simple enough. And it is worth noting that large, public-minded funders of HSR might be able to play a role here, for example requiring that grantees recognize and engage with subject organizations who seek to negotiate, and perhaps even helping to facilitate subject organizing where barriers are likely to stand in the way of organic development. These funders could potentially set a new standard for ethical research engagement, as they have in other areas.³⁴¹

* * *

337 A related issue is whether there would be any role for IRBs if subjects could collectively bargain. It is possible IRBs would still have an important function related to protecting community interests in sound science and the like, but it is unnecessary to substantially delve into this issue here given that collective bargaining is likely to occur so infrequently in the research setting.

338 GALL, *supra* note 75, at 225; Befort, *supra* note 312, at 361-77.

339 Note, however, that even in the standard employment context there is widespread agreement that these rights are not adequately protected. See, e.g., Benjamin I. Sachs, *Employment Law as Labor Law*, 29 CARDOZO L. REV. 2685, 2694-2700 (2008).

340 See, e.g., Domestic Workers United et al., *Domestic Workers and Collective Bargaining: A Proposal for Immediate Inclusion of Domestic Workers in the New York State Labor Relations Act*, NAT'L DOMESTIC WORKERS ALLIANCE 12-13 (Oct. 2010), <http://www.domesticworkers.org/sites/default/files/pdfs/collectivebargaining.pdf> (discussing potential ways to allow domestic workers to engage in collective bargaining, including how bargaining units could be defined, that may be of interest as an analogy to HSR).

341 See, e.g., *Guidance for Addressing the Provision of Antiretroviral Treatment for Trial Participants Following Their Completion of NIH-Funded HIV Antiretroviral Treatment Trials in Developing Countries*, NAT'L INSTS. HEALTH, <http://grants.nih.gov/grants/policy/antiretroviral/> (last updated June 9, 2005).

Considering the divergent interests of research subjects and those conducting the research, as well as issues of dependency and unequal bargaining power, features that generally motivate a variety of worker protections, there was already a strong case for leveling subjects up by revising the HSR regulations to incorporate certain protections from worklaw. This section strengthened that case in a number of areas by considering and rejecting possible reasons that worker protections would be inappropriate for research subjects.

Thus, with a few caveats regarding the continued ability to conduct essential research and certain inherent features of research injuries, the HSR regulations should permit subjects to be paid without regard to an upper limit, require that paid healthy subjects participating in nontherapeutic research be guaranteed some minimum wage, mandate that injured subjects be compensated for medical care and lost wages, extend consideration to site conditions, and protect subjects in their concerted activities and efforts at representative collective bargaining. Subjects need not be offered unemployment compensation, however, and ought to be encouraged to pursue more viable alternatives to unionization, such as the development of trade associations and other advocacy groups.

Where the regulations need revision, it clearly will not be appropriate to simply cut-and-paste from the relevant provisions protecting workers. Tailoring to account for the vagaries of HSR will be in order, but what has been established here is that those bioethicists advocating for worker protections in relatively abstract terms have been largely vindicated in the wake of rigorous legal and normative analysis.

V. OUTSTANDING OBJECTIONS

Before concluding, there are a handful of outstanding objections worthy of brief attention beyond those that have already been considered and dismissed above. First, as recognized at a number of points in the preceding analysis, worklaw as it currently exists is far from perfect – too many workers fall outside the scope of protection, minimum wages are too low, workers' compensation systems are flawed, union activity is not adequately supported, and US worklaws do not apply to workers at foreign sites.³⁴² Thus one may take issue with the use of worklaw as the appropriate lodestar for human subjects research regulation. To be clear, however, nothing herein is intended to suggest that if we just treat subjects like workers, no further action will be needed. Again, the fundamental point is recognizing the similarities between the two groups and the consistency

342 Zatz, *supra* note 81, at 58 n.3. On the question of international application, note that unlike worklaw, the U.S. HSR regulations do not simply defer to local standards, but rather apply equally whether research is conducted domestically or abroad (assuming the requisite jurisdictional link to the U.S.). See *Moral Science*, *supra* note 17, at 31-32. Thus, if the regulations are amended to add various worklaw protections as advocated herein, those protections would apply abroad as well. In this sense, international subjects would be "leveled beyond" their workplace counterparts.

in treatment that those similarities demand. Thus, if worker protections are inadequate for workers, they ought to be improved, and for subjects as well. But that fact is no reason not to extend the protections that do exist to subjects; subpar protections are preferable to none at all.

A related objection suggests that even if subjects are like workers, they should not be compared to workers in legally permissible industries but rather to those engaged in illegal work, or work that is prohibited to be paid, such as commercial sex work. Thus, the argument goes, even when it appears that subjects are more protected than workers, they are in fact less protected than the relevant comparators who are shielded by outright prohibition. The strongest response to this objection is that many of those other transactions should not be prohibited at all, but rather carefully regulated – just like HSR; indeed, a number of serious and convincing arguments have been made in favor of legalized or decriminalized prostitution, as well as permissible payment to organ “donors.”³⁴³ Moreover, the fact remains that we *do* allow subjects to be paid, and in that context, paid work is the most apt analogy.

Even those who accept the relevance of the comparison to workers, however, might be concerned that there is something fundamentally disrespectful about recasting the human research subject as human research worker, perhaps moving subjects from a special, revered category to the mundane ranks of the fast food employee or factory worker, for example. But note that the extension of various worker protections has been the result of the labor movement’s attempts over centuries to improve the worker’s plight and garner respect for the worker as a person rather than a widget. Thus, rather than treating the moniker “human research worker” as a demotion, it might even be considered a compliment.

Finally, to be absolutely clear, nothing herein is intended to suggest that repeat or professional research participation is a good thing. In most cases, it is decidedly not. So while some have expressed worry that extending greater protections to research subjects will make repeat participation more attractive, that fear can and should be addressed by the sorts of limitations on repeat participation discussed above, most importantly a subject registry.³⁴⁴ Even in the absence of a subject registry, however, it is important to recognize two things. First, repeat participation already occurs, and trying to discourage it via inadequate or inappropriate subject protections seems not only unfair, but callous and misguided. Second, the additional protections would apply not only to those subjects who consider research participation a job, but to all types of subjects, healthy and sick, paid and unpaid, one-time players and repeat enrollers. Thus, withholding protections to target one group of potential subjects is too blunt an instrument, and would ultimately leave many other subjects who are not engaged

343 See, e.g., Martha C. Nussbaum, “*Whether from Reason or Prejudice*”: Taking Money for Bodily Services, 27 J. LEG. STUD. 693 (1998).

344 It will also likely be necessary to take additional steps to make sure that adequate enrollment can be achieved with these limitations, or to simply accept the consequence that some research may not be possible without the inclusion of repeat participants.

in worrisome enrollment without valuable protections – protections that may make research participation itself more attractive to those who might otherwise have been unwilling to participate at all.

CONCLUSION

We often use the term “work” without really considering what it means – what should be included and excluded from its reach. There are areas of clear agreement: the miner is working, the nurse is working, the bus driver, the teacher, the firefighter. And there are areas of contention: stay-at-home parents, prostitutes, reality TV stars – and research subjects. But instead of focusing on whether a particular activity should be classified as work, a job, an occupation, a profession, or something else, it is fruitful to ask a different question: why do we extend certain legal protections to those who are engaged in work? And if those reasons also apply to other activities, even when those activities seem to fall outside the traditional boundaries of “work,” why should they be treated any differently? These are the questions that have driven this Article, and as a result, it is hoped that even those who reject the idea of research participation as work can see why certain worklaw protections ought to be extended to the research context.

Unlike the bioethicists who initiated this debate, this Article clarifies that a broad analogy between research subjects and workers does not actually suffice to demonstrate that subjects should be granted additional legal protections, since not all workers are protected by the law. And unlike the few legal decision-makers who have considered the employment status of research subjects, this Article goes beyond existing problematic legal distinctions to consider the ways in which subjects are like those workers who *should* be protected by the law. Ultimately, the HSR regulations should be revised such that all types of biomedical research subjects are free to accept unrestricted payment for their participation, eligible for a modified minimum wage, ensured no-fault compensation in the event of research injury, and protected in their efforts at concerted activity.

In the end, human subjects research may not be as special as it appears at first glance, nor the current regulatory scheme as protective as it might seem. The worker analogy draws these features into sharp relief, and should be pursued further, alongside other relevant comparisons, in order to develop the most consistent and justifiable legal and ethical approach to the acceptable involvement of human subjects in this socially important endeavor.

Criminal Law and HIV Testing: Empirical Analysis of How At-Risk Individuals Respond to the Law

Sun Goo Lee*

ABSTRACT:

This Note assesses the effect of laws that specifically criminalize behaviors that expose others to the human immunodeficiency virus (HIV). This Note examines the relationship between HIV testing decisions by high-risk individuals and the existence of these HIV-specific statutes, as well as the amount of media coverage related to them.

One of the main reasons public health experts criticize criminalization of HIV-exposing behavior is that it may discourage at-risk individuals from undergoing HIV testing. This argument, however, remains empirically untested to date. This study quantitatively examines whether at-risk individuals living in jurisdictions with HIV-specific statutes are less likely to report having been tested for HIV in the past year compared to those living in jurisdictions without HIV-specific statutes. Regression analysis is conducted using data collected in the United States over a seven-year span.¹

The results show that at-risk individuals residing in states with HIV-specific statutes are no less likely to report having been tested for HIV than those who live in other states. However, the number of people who reported that they had been tested for HIV is inversely correlated with the frequency of newspaper coverage of criminalization of HIV-exposing behavior. These findings imply that at-risk individuals' HIV testing is associated with media coverage of criminalizing HIV-exposing behavior.

The negative impact that criminal law has on HIV testing rates could be a serious public health threat. Testing is often the initial step in public health interventions that most effectively modify the risky behavior of HIV-positive individuals. The adverse consequence of criminalization should weigh heavily in the design and application of criminal sanctions for HIV-exposing behavior. In addition, future research should further explore the relationships between

* Judicial Researcher at the Supreme Court of Korea; SJD, 2013, Georgetown University Law Center; JSM, 2007, Stanford Law School; MPH, 2007, Seoul National University; LLB, 2004, Seoul National University. Since this note's acceptance into the Journal, the author has become Assistant Professor at Gachon University. The author is grateful to Professor Kathryn Zeiler and Professor Lawrence O. Gostin for their insightful comments.

¹ This study uses data from the Behavioral Risk Factor Surveillance System (BRFSS) collected by the Centers for Disease Control and Prevention (CDC) from 2002 to 2009 (with the exception of 2007) in 50 states and the District of Columbia in the United States.

criminalization, media coverage of criminalization, and HIV testing decisions for a more nuanced understanding of the consequences of criminalization.

TABLE OF CONTENTS

INTRODUCTION 198

**I. THE LAW AND ITS IMPACT ON PUBLIC HEALTH:
CRIMINALIZATION OF HIV-EXPOSING BEHAVIOR IN THE
UNITED STATES 199**

A. HISTORY OF CRIMINALIZATION OF HIV-EXPOSING BEHAVIOR IN THE
UNITED STATES..... 200

B. REVIEW OF HIV-SPECIFIC STATUTES IN 33 STATES 204

1. CRIMINAL INTENT 205

2. CRIMINAL BEHAVIOR 206

3. LACK OF DEFENSE: PRIOR DISCLOSURE OF HIV STATUS OR
PERFORMANCE OF SAFER SEX 208

C. THE GENERAL STD STATUTE OF KANSAS 208

D. REVIEW OF CASES APPLYING TRADITIONAL CRIMINAL LAW TO HIV-
EXPOSING BEHAVIOR..... 209

1. CRIME DISTINCTIONS UNDER HIV-SPECIFIC STATUTES..... 209

2. BITING AND SPITTING BY HIV-POSITIVE INDIVIDUALS..... 211

E. ARGUMENTS SURROUNDING CRIMINALIZATION OF HIV EXPOSURE AND
AVAILABLE EVIDENCE..... 212

1. ARGUMENTS ADVOCATING CRIMINALIZATION OF BEHAVIOR THAT
EXPOSES OTHERS TO HIV 213

2. CRITICISMS AGAINST CRIMINALIZATION AS A DISINCENTIVE FOR HIV
TESTING..... 215

**II. EMPIRICAL ANALYSIS OF HIV-SPECIFIC CRIMINAL STATUTES’
IMPACT ON HIV TESTING BEHAVIOR OF AT-RISK
INDIVIDUALS 218**

A. METHODOLOGIES 218

1. REGRESSION ANALYSIS..... 218

2. DESCRIPTION OF SAMPLES 230

B. RESULTS 233

C. DISCUSSION OF RESULTS..... 235

1. REGARDING HIV-SPECIFIC LAWS’ INFLUENCE ON HIV TESTING 235

2. THE MEDIA’S INTERMEDIARY ROLE BETWEEN CRIMINALIZATION AND HIV TESTING BEHAVIOR.....	236
III. GUIDANCE FOR FUTURE RESEARCH EFFORTS	237

INTRODUCTION

This Note assesses how criminalization of behavior that exposes others to HIV affects HIV testing decisions of at-risk individuals in the United States. Since the outbreak of the HIV epidemic in the 1980s, many state and federal courts have imposed criminal liability on individuals who expose others to HIV. Over half (33) of the jurisdictions in the United States have adopted HIV-specific statutes that impose criminal liability on HIV-positive individuals who expose others to HIV. Several others (six) have applied general criminal law (which I refer to as “traditional” law) and the state of Kansas has applied a statutory law that prohibits exposing others to a sexually transmitted disease (STD).² Estimates suggest that, to date, more than 900 HIV-positive individuals have been criminally prosecuted under these laws.³

This Note focuses on the impact of criminal law on public health policies, specifically in the context of HIV testing. Arguments that support or criticize criminalization center on the influence of criminal law on HIV-exposing and HIV testing behavior. Supporters of criminalization highlight criminal law’s deterrent effect on HIV-exposing behavior.⁴ Their primary claim is that the possibility of prosecution discourages HIV-positive individuals from engaging in risky sexual behaviors that spread the virus. Opponents claim that criminalization discourages HIV testing. According to this argument, the chance of criminal liability may discourage at-risk individuals from confirming their HIV status because an awareness of their status can assist the prosecution in establishing intent to infect others.⁵ This deterrence is a serious public health threat since HIV testing is crucial to HIV prevention.⁶

Several empirical studies have validated the theory that criminal punishment

2 As of 2012, 40 jurisdictions in the United States criminalize behavior that exposes others to HIV. Thirty-three states have adopted HIV-specific criminal statutes, six states have applied traditional criminal law to sanction individuals who expose others to HIV, and one state has applied a general STD statute regarding HIV-exposing behaviors. For detailed information, see Section I.A below.

3 *Global Criminalisation Scan: United States of America*, GNP+, <http://www.gnpplus.net/criminalisation/country/united-states-america> (last visited Dec. 15, 2013).

4 See *infra* Subsection I.E.2.

5 *Id.*

6 A number of studies have reported positive effects of public health interventions that reduce the incidence of new HIV infections. See, e.g., Reuben M. Granich et al., *Universal Voluntary HIV Testing with Immediate Antiretroviral Therapy as a Strategy for Elimination of HIV Transmission: A Mathematical Model*, 393 LANCET 48 (2009); Gary Marks et al., *Meta-Analysis of High-Risk Sexual Behavior in Persons Aware and Unaware They Are Infected with HIV in the United States*, 39 J. AIDS 446 (2005); Lance S. Weinhardt et al., *Effects of HIV Counseling and Testing on Sexual Risk Behavior: A Meta-Analytic Review of Published Research, 1985-1997*, 89 AM. J. PUB. HEALTH 1397 (1999).

fails to prevent risky sexual activities.⁷ However, researchers have not yet examined the impact of criminal law on HIV testing decisions.⁸ This Note fills this gap by analyzing whether and how HIV-specific criminal law affects HIV testing decisions. This study conducts regression analysis using the Behavioral Risk Factor Surveillance System (BRFSS) data that the Centers for Disease Control and Prevention (CDC) collected in the United States from 2002 to 2009 (except 2007). To contextualize its empirical findings, this Note starts with a normative analysis of HIV-criminalization laws.

The empirical study of criminal law in this Note suggests that current criminal laws are written and applied in a way that might be detrimental to HIV prevention policies. The regression results show that, in states with HIV-specific statutes, the number of at-risk individuals who report that they had been tested for HIV in the past 12 months is negatively correlated with the number of media reports on the criminalization of HIV-exposing behavior.

This Note's conclusions provide insight into how criminal law can interact with public health policy, and recommends that lawmakers consider the objectives of criminal and public health policies in tandem. The study's implications also apply to laws governing other communicable diseases, particularly those that share traits with HIV.

This Note proceeds as follows. Part I provides background information on criminal law governing HIV-exposing conduct. Section I.A explains the history of criminalizing HIV-exposing behavior in the United States. Sections I.B, C, and D respectively provide a normative analysis of HIV-specific statutes, traditional criminal law, and an STD statute that may apply to HIV-exposure cases. Section I.E presents a summary of two primary arguments surrounding the criminalization of behavior that exposes others to HIV, and discusses the validity of each argument based on available empirical evidence. Part II presents this study's methodologies and results, followed by the conclusion and suggestions for future research in Part III.

7 Scott Burris et al., *Do Criminal Laws Influence HIV Risk Behavior? An Empirical Trial*, 39 ARIZ. ST. L.J. 467 (2007); Catherine Dodds et al., *Responses to Criminal Prosecutions for HIV Transmission Among Gay Men with HIV in England and Wales*, 17 REPROD. HEALTH MATTERS 135 (2009); Keith J. Horvath et al., *Should It Be Illegal for HIV-Positive Persons to Have Unprotected Sex Without Disclosure? An Examination of Attitudes Among US Men Who Have Sex with Men and the Impact of State Law*, 22 AIDS CARE 1221 (2010). For a detailed summary of these studies, see Subsection I.E.1 below.

8 The SERO Project performed a study on this subject, but has published neither methodologies nor final results yet. For preliminary findings, see *The SERO Project: National Criminalization Preliminary Results*, SERO (July 25, 2012), http://seroproject.com/wp-content/uploads/2012/07/Sero-Preliminary-Data-Report_Final.pdf.

I. THE LAW AND ITS IMPACT ON PUBLIC HEALTH: CRIMINALIZATION OF HIV-EXPOSING BEHAVIOR IN THE UNITED STATES

A. History of Criminalization of HIV-Exposing Behavior in the United States

The majority (40) of United States jurisdictions penalize individuals who expose others to HIV by applying an HIV-specific criminal statute, a general STD statute, or a traditional criminal law (Table 1).⁹ Thirty-three states have introduced HIV-specific statutes to penalize HIV-exposing behavior. One state, Kansas, has applied a general STD statute that applies to HIV-positive individuals who expose others to the virus. The courts of six states have been applying traditional crimes, such as attempted murder and aggravated assault, to individuals who expose others to HIV.¹⁰ As a consequence, people in these states, who account for more than 90% of the United States population, live under legal systems that criminally punish people with HIV who expose others.¹¹

Table 1. Summary of State Legislation Against HIV Exposure

Thirty-three states with HIV-specific criminal laws	Alabama, Alaska, Arkansas, California, Colorado, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maryland, Michigan, Mississippi, Missouri, Montana, Nevada, New Jersey, North Carolina, North Dakota, Ohio, Oklahoma, Pennsylvania, South Carolina, South Dakota, Tennessee, Utah, Virginia, Washington, Wisconsin
One state applying general STD statutes	Kansas
Six states applying traditional criminal law	Massachusetts, Minnesota, New Hampshire, New York, Oregon, Texas

9 State laws were collected from three different sources: (1) *Global Criminalisation Scan*, *supra* note 3; (2) *HIV Criminalization: State Laws Criminalizing Conduct Based on HIV Status*, LAMBDA LEGAL, http://www.lambdalegal.org/publications/fs_hiv-criminalization (last updated July 12, 2010); and (3) *State Criminal Statutes on HIV Transmission—2008*, AM. CIVIL LIBERTIES UNION (ACLU) (2008), http://www.aclu.org/files/images/asset_upload_file292_35655.pdf. Each source's information was cross-checked with other sources to obtain the most up-to-date survey of each state's law. When there was a discrepancy between the reports of three organizations, Westlaw and LexisNexis search engines were used to obtain the texts of the statute and verify the criminalizing policy of each state.

10 *HIV Criminalization Fact Sheet*, CTR. FOR HIV LAW & POLICY (Nov. 14, 2012), http://www.hivlawandpolicy.org/sites/www.hivlawandpolicy.org/files/PJP%20fact%20sheet_11.14.2012.pdf.

11 As of 2010, 308,745,538 people live in the United States, 287,508,591 (93.12%) of whom reside in the 40 states that criminalize behavior that exposes others to HIV. U.S. CENSUS 2010, <http://www.census.gov/2010census/> (last visited Nov 11, 2013).

**Eleven states *not*
criminalizing HIV exposure**

Arizona, Connecticut, Columbia, Hawaii, Maine, Nebraska, New Mexico, Rhode Island, Vermont, West Virginia, Wyoming

HIV-specific legislation began in Washington in 1988.¹² At the onset of the epidemic, the Presidential Commission on the Human Immunodeficiency Virus Epidemic submitted a report to the President, which recommended prosecuting HIV-positive individuals who subject others to a risk of infection.¹³ The Report explained that criminalization holds violators of the law accountable for their conduct and also deters high-risk behavior.¹⁴ The Ryan White Comprehensive AIDS Resources Emergency Act (the CARE Act) of 1990 also catalyzed HIV-specific criminal legislation.¹⁵ This Act restricts federal emergency AIDS relief grants to jurisdictions with laws that criminalize HIV-exposing behaviors.¹⁶ Jurisdictions can fulfill this requirement by amending their public health statutes to include HIV under their existing STD exposure statutes, applying traditional criminal law such as attempted murder or aggravated assault, or introducing an HIV-specific criminal statute.¹⁷ Following enactment of the CARE Act, jurisdictions that did not already have HIV-specific statutes added them in haste. By 1990, 21 states had introduced statutes that penalize behavior that exposes others to HIV through sexual means.¹⁸ Jurisdictions that already had statutes amended them to comply with the Act's requirements.¹⁹

The second wave of HIV-specific state legislation occurred around 1998, in response to the Nushawn Williams case. In 1997, Williams was prosecuted for having unprotected sex with approximately 50 women in New York after learning he was HIV-positive.²⁰ The media covered the story heavily, leading to widespread anger against the HIV-positive population.²¹ The Williams case prompted states to adopt HIV-specific criminal statutes in order to remove

12 Arianne Stein, *Should HIV Be Jailed? HIV Criminal Exposure Statutes and Their Effects in the United States and South Africa*, 3 WASH. U. GLOBAL STUD. L. REV. 177, 181 (2004).

13 REPORT OF THE PRESIDENTIAL COMMISSION ON THE HUMAN IMMUNODEFICIENCY VIRUS EPIDEMIC 130, 152 (1988) [hereinafter PRESIDENTIAL COMMISSION REPORT].

14 *Id.*

15 Jodi Mosiello, *Why the Intentional Sexual Transmission of Human Immunodeficiency Virus (HIV) Should Be Criminalized through the Use of Specific HIV Criminal Statutes*, 15 N.Y.L. SCH. J. HUM. RTS. 595, 598 (1999).

16 Ryan White Comprehensive AIDS Resources Emergency Act of 1990, Pub. L. No. 101-381, § 301(a), 104 Stat. 576, 603 (repealed 2000).

17 Stephen V. Kenney, Comment, *Criminalizing HIV Transmission: Lessons from History and a Model for the Future*, 8 J. CONTEMP. HEALTH L. & POL'Y 245, 247 (1992).

18 *Id.* at 840.

19 Amanda Weiss, Comment, *Criminalizing Consensual Transmission of HIV*, 2006 U. CHI. LEGAL F. 389, 392 (2006); see also Leslie E. Wolf & Richard Vezina, *Crime and Punishment: Is There a Role for Criminal Law in HIV Prevention Policy?*, 25 WHITTIER L. REV. 821, 844 (2004).

20 Stein, *supra* note 12, at 180.

21 As one commentator noted, such "highly publicized, outrageous cases" caused a "public uproar." *Id.*

barriers against prosecuting HIV-exposing individuals.²² In Florida, for example, state legislators cited the Williams case when they classified knowing transmission (or attempted transmission) of HIV as a felony.²³ A statement by an Ohio lawmaker when he introduced a criminalization bill characterizes public sentiment at the time: “It is wrong for society to simply look the other way and not offer reasonable protection to those who are unknowingly being exposed to this lethal disease.”²⁴

Along with HIV-specific statutes, some jurisdictions have applied traditional crimes to prosecute individuals who knowingly expose others to HIV. Some states that adopted HIV-specific statutes at a later stage of the epidemic had previously used traditional criminal laws to prosecute HIV-positive individuals. States without HIV-specific legislation often issue charges based on the traditional crimes of attempted murder and aggravated assault.²⁵

According to data released by the Global Network of People living with HIV/AIDS (GNP+), there have been over 900 incidents of arrest or prosecution of HIV-positive individuals who allegedly exposed others to HIV in the United States.²⁶ Table 2 provides the year each state started criminalizing behavior that exposes others to HIV, the accumulated number of prosecutions and convictions reported for each state court, and the type of crime exposing others to HIV would constitute in that state.

Table 2. Overview of Criminalization of HIV Exposure in the United States

State	Year of Criminalization	Number of Prosecutions	Number of Convictions	Type of Crime
Alabama	1987	1	0	Class C misdemeanor
Alaska	1996 ²⁷	0	0	Sentence enhancement
Arkansas	1989	8+	5+	Class A felony
California	1998	10+	Unknown	Felony

²² *Id.*

²³ THOMAS C. SHEVORY, NOTORIOUS HIV: THE MEDIA SPECTACLE OF NUSHAWN WILLIAMS (2004).

²⁴ Mark Tatge, *Bill Would Require HIV Disclosure*, PLAIN DEALER (Cleveland, Ohio), Feb. 10, 1999, at 5B.

²⁵ For criminal charges brought under traditional criminal laws, see *infra* Section I.D.

²⁶ National AIDS Manual (NAM) also estimates that at least 442 HIV-positive individuals have been prosecuted, and emphasizes that “it is likely that [this] estimate . . . substantially underestimates the actual number.” *HIV & the Criminal Law: North America*, NAM AIDSMAP, <http://www.aidsmap.com/page/1445031> (last visited Dec. 18, 2013).

²⁷ Alaska first applied traditional criminal law to an HIV exposure case in 1996, see *Wans v. State*, No. A-6188, 1996 WL 671355 (Alaska Ct. App. Nov. 20, 1996), and then adopted an HIV-specific statute in 2006.

CRIMINAL LAW AND HIV TESTING

Colorado	1999	6+	4+	Sentence enhancement
Delaware	1988	0	0	Class E felony
Florida	1997	239	Unknown	3rd degree felony
Georgia	1988	20+	10+	Felony
Idaho	1988	54	Unknown	Felony
Indiana	1988	20+	15+	Class A/C felony
Illinois	1989	100	Unknown	Class 2 felony
Iowa	1998	25	15	Class B felony
Kansas	2009	1	1	Level 7 person Felony
Kentucky	1990	8+	Unknown	Class D felony
Louisiana	1987	55+	Unknown	Not specified ²⁸
Maryland	1989	8+	5+	Misdemeanor
Massachusetts	2003 ²⁹	4+	1+	General criminal law
Michigan	1989	Unknown	5+	Felony
Minnesota	1988 ³⁰	3+	2+	General criminal law
Mississippi	1999 ³¹	2	2	Felony
Missouri	1988	27+	10+	Class A/B felony
Montana	1989	0	0	Misdemeanor
Nevada	1993	3+	3+	Category B felony
New Hampshire	2002 ³²	2	1	Class B felony
New Jersey	1997	4	3	Crime of 3 rd degree

28 Punishable by a fine not more than \$6000, imprisonment with or without hard labor for not more than one year, or both, except in certain circumstances.

29 Commonwealth v. Smith, 790 N.E.2d 708, 709 (Mass. App. Ct. 2003).

30 U.S. v. Moore, 846 F.2d 1163, 1164 (8th Cir. 1988).

31 Mississippi first punished HIV exposure through traditional criminal law in 1999. Carter v. State, 803 So. 2d 1191 (Miss. Ct. App. 1999). The state later introduced an HIV-specific statute in 2004.

32 State v. C.J., No. 01-S-726, 2002 WL 31059244 (N.H. Super. Ct. May 23, 2002).

New York	1988 ³³	6+	Unknown	Class D felony
N. Carolina	1988 ³⁴	6+	2+	Misdemeanor
North Dakota	1989	1	1	Class A felony
Ohio	1996 ³⁵	33+	19+	2 nd degree felony
Oklahoma	1988	6+	Unknown	Felony
Oregon	1992	3+	Unknown	Felony
Pennsylvania	1998	13+	Unknown	2 nd degree felony
S. Carolina	1988	11+	5+	Felony
South Dakota	2000	69	Unknown	Class 3 felony
Tennessee	1994	50+	40+	Class C felony
Texas	1998	22+	20+	2 nd degree felony
Utah	1993	1+	1+	Sentence enhancement
Virginia	1995 ³⁶	3+	Unknown	Class 6 felony
Washington	1997	9+	6+	Class A felony
Wisconsin	2001	0	0	Sentence enhancement
Total		343+	206+	

B. Review of HIV-Specific Statutes in 33 States

In general, HIV-specific statutes require three elements to establish criminal HIV exposure: criminal intent, criminal behavior, and lack of defense.³⁷ This section reviews how courts of different jurisdictions define each of these three elements.

33 *People v. Hawkrigg*, 525 N.Y.S. 2d 752, 753 (Suffolk Cnty. Ct. N.Y. 1988).

34 10A N.C. ADMIN. CODE 41A.0202 (1988).

35 Ohio first applied traditional criminal law in 1996. *State v. Bird*, No. 96APA04-505, 1996 WL 751467 (Ohio Ct. App. Dec. 31, 1996), *aff'd*, 692 N.E.2d 1013 (Ohio 1998). Ohio adopted an HIV-specific statute later in 2000.

36 Traditional criminal law was first applied to an HIV exposure case in Virginia in 1995. *United States v. Sturgis*, 48 F.3d 784 (4th Cir. 1995). Virginia introduced an HIV-specific statute in 2000.

37 Stein, *supra* note 12, at 182.

1. Criminal Intent

In general, statutory crimes require that a defendant have a certain state of mind—*mens rea*—often defined as “intentionally,” “knowingly,” “purposely,” or “willfully.”³⁸ The basic idea is that “an act does not make [a person] guilty, unless the mind be guilty.”³⁹ HIV-specific criminal laws also require that an HIV-positive individual has *mens rea*. The laws of six states—California, Kansas, Oklahoma, South Dakota, Virginia, and Washington—require that a defendant have the specific intent to transmit or expose HIV to others to bear criminal liability.⁴⁰ The remaining state laws (28) simply require that individuals are aware of their HIV-positive status.⁴¹ Among these 28 states, the laws of five define when an HIV-positive individual has knowledge of HIV infection,⁴² but the other 23 are silent on this point. As a result, whether an individual who has received a medical diagnosis of AIDS has sufficient knowledge of HIV infection or an HIV positive test result is necessary is largely left to judicial interpretation.

Except in Florida and Kentucky, HIV-specific criminal statutes of 33 states do not address whether HIV-positive individuals should understand that they are

38 WAYNE R. LAFAVE, CRIMINAL LAW, 239-41 (4th ed., 2003).

39 JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 117 (5th ed. 2009); LAFAVE, *supra* note 38, at 239.

40 The statute of California clearly states that a person’s “knowledge of his or her HIV-positive status, without additional evidence, shall not be sufficient to prove specific intent.” CAL. HEALTH AND SAFETY CODE § 120291 (West 2013). “It is unlawful for an individual . . . with the intent to expose that individual to that life threatening communicable disease.” KAN. STAT. ANN. § 21-3435 (2013). In Oklahoma, it is unlawful for an HIV-positive individual to engage in certain conduct “with intent to infect another.” OKLA. STAT. tit. 21 § 1192.1 (2013). But when HIV-positive status enhances the penalty for other sex crimes such as prostitution, laws generally do not require an HIV-positive individual to have intent.

41 “Knowingly” or “with knowledge”: ALA. CODE § 22-11A-21(c) (2013); COLO. REV. STAT. §§ 18-3-415.5, -7-205.7 (2013); GA. CODE ANN. § 16-5-60(C) (2013); IDAHO CODE ANN. § 39-608 (West 2013); ILL. COMP. STAT. § 5.12-16.2 (2013); IOWA CODE § 709C.1 (2013); MD. CODE ANN., HEALTH-GEN. § 18-601.1 (West 2013); MICH. COMP. LAWS ANN. § 333.5210 (2013); MINN. STAT. § 609.2241 (2013); MISS. CODE ANN. § 97-27-14(1) (2013); MONT. CODE ANN. §§ 50-18-112, -18-113 (2013); N.J. STAT. ANN. § 2C:34-5 (2013); OHIO REV. CODE ANN. § 2903.11 (2013); S.C. CODE ANN. § 44-29-145 (2013); TENN. CODE ANN. § 39-13-109 (2013); UTAH CODE ANN. § 76-10-1309 (2013); W. VA. CODE §§ 16-4-20, -4-26 (2013). “Intentionally, recklessly, or knowingly exposed others to HIV”: IND. CODE §§ 35-42-1-7, -2-6(e) (2013); NEV. REV. STAT. § 201.205 (2013); 18 PA. CONS. STAT. ANN. § 2703 (2013); S.D. CODIFIED LAWS §§ 22-18-31, -33, -34 (2013). “Willfully”: N.D. CENT. CODE § 12.1-20-17 (2013).

42 The laws of these five states are as follows: ALASKA STAT. ANN. § 12.55.155(c)(33) (West 2013) (stating that those “previously diagnosed as having or having tested positive for HIV” are criminally culpable); ARK. CODE ANN. § 5-14-123 (2013) (clearly stating that only an individual who had received a positive test is forbidden from exposing others to HIV); COLO. REV. STAT. § 18-7-201(7) (2013) (“with knowledge of being infected with” HIV); OHIO REV. CODE ANN. § 2903.11 (2013) (“with knowledge that the person has tested positive” for HIV); UTAH CODE ANN. § 76-10-1309 (2013) (“[a] person . . . is guilty of a third degree felony if the person is an HIV positive individual”).

infectious.⁴³ There is a difference between simply knowing one's HIV-positive status and understanding the risk of infecting others, but most states only require the former to establish criminal intent for the crime of exposing others to HIV.

2. Criminal Behavior

Many state laws provide only a very general and blurry depiction of what behavior the law proscribes. In Alabama, for example, a statute forbids "*any act which will probably or likely transmit such disease to another person.*"⁴⁴ Mississippi forbids an HIV-positive individual from "*expos[ing] another person to HIV,*"⁴⁵ but it does not define exposure. The law of Illinois punishes HIV-positive individuals if they knowingly "engage in intimate contact with another" or "transfer . . . blood, tissue, semen, organs, or other *potentially infectious* body fluids . . . to another."⁴⁶ The term "potentially infectious body fluids" is quite broad and could be interpreted to include tears or saliva, which have a negligible risk of HIV transmission.⁴⁷ The law also defines "intimate contact" vaguely, as "the exposure of the body of one person to a bodily fluid of another person in a

43 FLA. STAT. ANN. § 384.24(2) (West 2013) (with knowledge of such infection and having "been informed that he or she may communicate this disease to another person through sexual intercourse"); KY. REV. STAT. § 311.990(24)(b) (West 2013) ("knows he is infected with human immunodeficiency virus, and who has been informed that he may communicate the infection by donating organs, skin, or other human tissue").

44 ALA. CODE § 22-11A-21(c) (West 2013) (emphasis added); *see also* NEV. REV. STAT. § 201.205 (West 2013) ("engages in conduct in a manner that is intended or likely to transmit the disease to another person").

45 ALASKA STAT. ANN. § 12.55.155(c)(33) (West 2013) ("exposed the victim to a risk or a fear that the offense could result in the transmission of HIV"); MISS. CODE ANN. § 97-27-14(1) (West 2013) ("expose another person to [HIV]"); MONT. CODE ANN. §§ 50-18-112 (West 2013) ("expose another person to infection"); MO. REV. STAT. § 191.677 (West 2013) ("any individual knowingly infected with HIV . . . to [a]ct in a reckless manner by exposing another person to HIV").

46 720 ILL. COMP. STAT. ANN. 5/12-5.01 (West 2013) (emphasis added); *see also* Christina M. Shriver, *State Approaches to Criminalizing the Exposure of HIV: Problems in Statutory Construction, Constitutionality and Implications*, 21 N. ILL. U. L. REV. 319, 332 (2001).

47 Shriver, *supra* note 46, at 333. Other states have similar statutes in place: 720 ILL. COMP. STAT. ANN. 5/12-5.01 ("(1) engages in sexual activity with another without the use of a condom knowing that he or she is infected with HIV; (2) transfers, donates or provides his or her blood, tissue, semen, organs or other potentially infectious body fluids for transfusion, transplantation, insemination, or other administration to another"); IOWA CODE § 709C.1 (West 2013) ("(1) engages in intimate contact with another person; (2) transfers, donates, or provides blood, tissue, semen, organs, or other potentially infectious bodily fluids for transfusion, transplantation, insemination, or other administration to another person"); S.D. CODIFIED LAWS §§ 22-18-31, (West 2013) ("(1) engaging in sexual intercourse or other intimate physical contact with another person; (2) transferring, donating, or providing blood, tissue, semen, organs or other potentially infectious body fluids or parts for . . . administration to another in any manner that presents a significant risk of HIV infection"); TENN. CODE ANN. § 39-13-109 (West 2013) ("(1) [e]ngages in intimate contact with another; (2) [t]ransfers, donates, or provides blood, tissue, semen, organs, or other potentially infectious body fluids or parts for transfusion, transplantation, insemination, or other administration to another in any manner that presents a significant risk of HIV . . . transmission").

manner that *could* result in the transmission of HIV.”⁴⁸

The laws of many jurisdictions explicitly prohibit acts that have a negligible risk of HIV transmission. Louisiana, Pennsylvania, and Missouri criminalize biting and/or spitting by HIV-positive individuals as a form of HIV exposure,⁴⁹ although “transmission by biting is extremely rare.”⁵⁰ Even in these rare cases, the people who were bitten acquired HIV because their deep wounds were exposed to a substantial amount of blood in the biter’s saliva.⁵¹ Some state statutes, such as that of Illinois, prohibit exposure to bodily fluids such as urine and saliva, which do not have any reported incidents of HIV transmission.⁵² In Arkansas and Michigan, the definition of proscribed sexual activities includes intrusion of “any object,”⁵³ enabling prosecution of many activities that carry no risk of transmission. This broad definition may even prohibit the use of sex toys by HIV-infected people.⁵⁴

The majority of the aforementioned 33 states that have HIV-specific criminal statutes prohibit conduct even if it has only a slight risk of HIV transmission and do not require that an HIV-negative individual actually contract HIV as a result of the exposing incident.⁵⁵ Most of the states do not even require that the victim be HIV-free at the time of exposure. In fact, one of the reasons state legislatures adopted HIV-specific criminal statutes was to eliminate prosecutors’ burden of proving that the incident in question caused HIV infection. Under traditional

48 Shriver, *supra* note 46, at 332 (emphasis added). Iowa defines intimate contact as “the intentional exposure of the body of one person to a bodily fluid of another person in a manner that could result in the transmission of the human immunodeficiency virus.” IOWA CODE § 709C.1(2)(b) (2013). In Tennessee, the law defines an “intimate contact with another” as the “exposure of the body of one person to a bodily fluid of another person in any manner that presents a significant risk” of HIV transmission. TENN. CODE ANN. § 39-13-109(b)(2) (West 2013).

49 LA. REV. STAT. ANN. § 14:43.5 (West 2013); 18 PA. CONS. STAT. ANN. § 2703 (West 2013); MO. REV. STAT. § 191.677 (West 2013).

50 *HIV & the Criminal Law: Biting*, NAM AIDSMAP, <http://www.aidsmap.com/Biting/page/1322751> (last visited Dec. 18, 2013).

51 *Id.*

52 See Shriver, *supra* note 46, at 333-34; see also Carol L. Galletly & Steven D. Pinkerton, *Toward Rational Criminal HIV Exposure Laws*, 32 J.L. MED. & ETHICS 327, 329-31 (2004). Other states have similar laws that punish HIV-positive individuals for exposing others to bodily fluids, including saliva or urine. See, e.g., GA. CODE ANN. CODE § 16-5-60(d) (2013); IDAHO CODE § 39-608 (2013); 18 PA. CONS. STAT. ANN. § 2703.

53 See ARK. CODE ANN. § 5-14-123 (making it a class A felony to knowingly engage in any “intrusion, however slight, of any part of a person’s body or of any object into the genital or anal opening of another person’s body, without first having informed the other person of the presence of HIV”); MICH. COMP. LAWS ANN. § 333.5210 (2013) (the prohibiting of “any other intrusion, however slight, of any part of a person’s body or of any object into the genital or anal openings of another person’s body”).

54 Galletly & Pinkerton, *supra* note 52, at 329. The irony of these prohibitions is that the use of non-shared sex toys can be a satisfying risk-free alternative to intercourse with an HIV-infected sex partner. *Id.*

55 Andrew M. Francis & Hugo M. Mialon, *The Optimal Penalty for Sexually Transmitting HIV*, 10 AM. L. & ECON. REV. 388, 396 (2008).

criminal law, successful prosecution of murder or assault would require proven causation between an exposing act and HIV infection. However, because HIV infection has an initial asymptomatic period, which can last up to 10 years,⁵⁶ prosecutors often have difficulty establishing causation. HIV-specific statutes remove this prosecutorial burden.

3. Lack of Defense: Prior Disclosure of HIV Status or Performance of Safer Sex

HIV-specific statutes in many states establish prior disclosure of HIV-positive status as a legal defense.⁵⁷ In these states, if the defendant can prove that he or she informed a sex partner that he or she carried HIV, the defendant does not bear criminal liability. In some states, including California and North Dakota, the law requires that the infected individual take precautions in addition to disclosing infection. In California, a person is liable for exposing others to HIV only when the person has “not disclosed his or her HIV-positive status.”⁵⁸ In North Dakota, that intercourse took place after full disclosure of the risk of such activity and involved the use of an “appropriate prophylactic device” can serve as an affirmative defense.⁵⁹

In the minority of jurisdictions, such as Kansas, Maryland, and Montana, laws do not allow any of these defenses.⁶⁰ In these states, an HIV-infected individual could be criminally liable for intercourse even if he or she informs partners of his or her status and takes precautions to prevent infection during intercourse.

C. The General STD Statute of Kansas

Kansas has a general sexually transmitted disease (STD) statute that prohibits an infected individual from knowingly engaging in “sexual intercourse” with the intent of exposing a partner to the disease.⁶¹ The definition of the term “sexual intercourse” does not include penetration by any object other than the male sex organ.⁶²

In 2009, Kansas first applied this law to an HIV-positive individual who was accused of exposing his sex partners to HIV.⁶³ The defendant had both protected

56 *HIV/AIDS: HIV Basics*, CDC, <http://www.cdc.gov/hiv/basics/> (last updated Nov. 8, 2013).

57 For the texts of state statutes, see *HIV Criminalization: State Laws Criminalizing Conduct Based on HIV Status*, *supra* note 9.

58 CAL. HEALTH AND SAFETY CODE § 120291 (West 2013).

59 N.D. CENT. CODE § 12.1-20-17(3) (2013).

60 KAN. STAT. ANN. § 21-3435 (2013); MD. CODE ANN., HEALTH-GENERAL § 18-601 (2013); MONT. CODE ANN. § 50-18-112 (2013).

61 KAN. STAT. ANN. § 21-3435.

62 *Id.*

63 *Global Criminalisation Scan: USA—Kansas*, GNP+, <http://www.gnpplus.net/criminalisation/country/usa-kansas> (last updated Oct. 11, 2012).

and unprotected sexual intercourse with two women without first disclosing that he was HIV-positive. The Supreme Court of Kansas held that to establish a violation of the law,⁶⁴ the state must prove beyond a reasonable doubt that the defendant (1) knew that he was “infected with a life-threatening communicable disease”, (2) “knowingly engaged in sexual intercourse” with the victims, and (3) “engaged in this conduct with the intent of exposing” the victims to the disease.⁶⁵

The Supreme Court of Kansas held that exposing others to an STD is a “specific intent” crime in that the law “specifically identifies or requires the further particular intent to expose.”⁶⁶ The court stated that such specific intent could be inferred from the fact that the defendant knew he had HIV at the time of intercourse, none of the victims knew, he did not use a condom, and he falsely represented to a victim that he was free of HIV.⁶⁷ But in this case, the court decided that the prosecution failed to prove these points and therefore failed to show that the defendant had the specific intent to expose his partners. Accordingly, the Supreme Court overturned the defendant’s conviction.

D. Review of Cases Applying Traditional Criminal Law to HIV-Exposing Behavior

In Massachusetts, Minnesota, New Hampshire, New York, Oregon, and Texas, courts have applied traditional criminal laws to HIV-positive individuals who expose others to HIV.⁶⁸ The three types of behavior that are most frequently discussed in cases are sexual conduct, biting, and spitting on others.⁶⁹ Each type of behavior can result in charges for attempted murder and aggravated assault.

1. Crime Distinctions Under HIV-Specific Statutes

The most frequent criminal charge in sexual transmission cases is attempted murder.⁷⁰ Murder is often defined as the “unlawful killing of another ‘living

64 *State v. Richardson*, 209 P.3d 696, 703 (Kan. 2009).

65 *Id.*

66 *Id.* at 698.

67 *Id.* at 704.

68 Other states with HIV-specific criminal statutes also prosecuted behavior that exposes others to HIV based on traditional criminal law before they adopted HIV-specific statutes, but this section reviews how courts have shaped criminal liability for exposing others to HIV in states that only use traditional criminal law to penalize HIV-exposing behavior.

69 See *Global Criminalisation Scan: United States of America*, *supra* note 3.

70 Prosecution of an HIV-exposing individual for murder is highly unlikely to be successful. First, the prosecution must show that the defendant behaved “purposefully, knowingly, or recklessly with extreme indifference to the value of human life.” Rebecca Ruby, *Apprehending the Weapon Within: The Case for Criminalizing the Intentional Transmission of HIV*, 36 AM. CRIM. L. REV. 313, 325 (1999). That is, the prosecution must show that the defendant “knew or thought that he had HIV, that he subjectively believed the virus could be transmitted through his conduct, and that he wanted his behavior to lead to the death of his victim.” *Id.* at 326. This intent may be impossible to prove. Second, the prosecution is not likely to overcome the barrier of the causation

human being' with 'malice aforethought.'"⁷¹ "Attempted murder" is an act done with the intent to commit murder but falls short of its actual commission.⁷² In order to successfully prosecute an HIV-positive individual for attempted murder, the state has to prove that the infected individual had sexual intercourse with the specific intent to commit murder, even though the exposed individual did not die as a result of that encounter.⁷³

In cases where HIV exposure was charged as attempted murder, courts have focused on determining whether the defendant had the necessary criminal intent. In general, courts have held that individuals have the specific intent to commit murder when they have intercourse knowing that they are HIV-positive. In such cases, the prosecution need not prove that the defendant intended to harm or cause the death of the victim. In *State v. Hinkhouse*, for example, an HIV-positive defendant was accused of attempted murder for having unprotected sexual intercourse with various women without disclosing his HIV-positive status even after he found out that he was HIV-positive.⁷⁴ The defendant claimed that he meant only to satisfy himself sexually and that there was insufficient evidence to prove his intent to harm or to cause the death of others. The court, however, stated that having unprotected sex without disclosing his HIV-positive status "demonstrated that his purpose was more than mere sexual gratification."⁷⁵

As long as HIV-positive individuals considered death to be a possible consequence of exposing others to HIV, courts have found specific intent even when no death has resulted. In *Hinkhouse*, the defendant was convicted of ten counts of attempted murder and ten counts of attempted assault. The defendant appealed but the Court of Appeals of Oregon affirmed the lower court's decision, holding that the crime of attempted murder only requires criminal intent, not the consequence of death. According to the court, expert testimony that HIV is rarely transmitted through unprotected sexual intercourse does not affect criminal liability for attempted murder when the defendant engaged in sexual activities knowing that he tested positive for HIV and that exposing others to his HIV could result in HIV transmission.⁷⁶

requirement. Common law requires "that the victim die within a year and a day of the defendant's actions in order to attribute the death to the defendant." This requirement would preclude most murder prosecution for HIV because HIV has a lengthy incubation period and can progress slowly. *Id.*; see also LAWRENCE O. GOSTIN, *THE AIDS PANDEMIC: COMPLACENCY, INJUSTICE, AND UNFULFILLED EXPECTATIONS* 187-88 (2004).

71 LAFAVE, *supra* note 38, at 725.

72 DRESSLER, *supra* note 39, at 508-09; LAFAVE, *supra* note 38, at 725.

73 175A AM. JUR. 2D *Homicide* § 580 (2012).

74 *State v. Hinkhouse*, 912 P.2d 921 (Or. Ct. App. 1996).

75 *Id.* at 925.

76 In this case, the court stated that according to Oregon law, a person commits a crime "when the person intentionally engages in conduct which constitutes a substantial step towards commission of the crime." *Hinkhouse*, 912 P.2d. at 924 (quoting OR. REV. STAT. § 163.185(1) (1995)). A person, therefore, "commits attempted murder when he or she attempts, without justification or excuse, intentionally to cause the death of another human being." *Id.*

Individuals with HIV have also been prosecuted for aggravated assault, assault with a deadly weapon, or reckless conduct with a deadly weapon for having sex with an unknowing partner.⁷⁷ Courts have ruled that having unprotected intercourse without disclosing HIV-positive status meets the conduct standards required for these crimes; many have repeatedly held that seminal fluid containing HIV is a deadly weapon under state criminal laws.⁷⁸

Courts have further held HIV-positive individuals liable for aggravated assault regardless of whether the HIV-positive individuals actually put the victim in danger of contracting HIV. In *Lewis v. State*,⁷⁹ the defendant had inserted his finger into the vagina of a 10-year old girl and masturbated until he ejaculated near her. The Court of Appeals of Texas affirmed the trial court's judgment, convicting Lewis of aggravated sexual assault.⁸⁰ The Court stated that, despite the lack of actual transmission of body fluids that are likely to infect the victim with HIV, the aggravated crime applies to the defendant because he had HIV.⁸¹

2. Biting and Spitting by HIV-Positive Individuals

In addition to sexual activities, biting and spitting have led to charges of attempted murder against HIV-positive individuals. In these cases, courts have inferred intent to murder from what an HIV-positive individuals said when they bit or spit at others.⁸² In *State v. Smith*, an HIV-positive jail inmate was prosecuted for attempted murder after he bit an officer.⁸³ The defendant was convicted of attempted murder and aggravated assault before the superior court, and he appealed. The defendant argued that he did not have intent to kill because

77 "A person is guilty of aggravated assault if he: (a) attempts to cause serious bodily injury to another, or causes such injury purposely, knowingly or recklessly under circumstances manifesting extreme indifference to the value of human life; or (b) attempts to cause or purposely or knowingly causes bodily injury to another with a deadly weapon." MODEL PENAL CODE § 211.1(2) (1962). Assault with a deadly weapon (alternatively referred to as "assault with a dangerous weapon") is defined as "[a]n aggravated assault in which the defendant, using a deadly weapon, threatens the victim with death or serious bodily injury." BLACK'S LAW DICTIONARY (9th ed. 2009). See also *Sellers v. State*, No. 05-94-0033-CR, 1996 WL 223537 (Tex. Ct. App. Apr. 29, 1996); *Hinkhouse*, 912 P.2d at 925.

78 *State v. C.J.*, No. 01-S-726, 2002 WL 31059244 (N.H. Super. Ct. May 23, 2002); *Sellers*, 1996 WL 223537; *Hinkhouse*, 912 P.2d at 924-25.

79 *Lewis v. State*, No. 07-08-00290-CR, 2010 WL 2400085 (Tex. Ct. App. June 16, 2010).

80 *Id.*

81 *Id.* In this case, the court also mentioned that, although the defendant merely ejaculated, the fact that there was a semen stain on the shorts he was wearing made it a *possibility* that the victim might have been exposed to HIV because she had "a light brown discharge from the vaginal area" from rubbing with defendant's finger. *Id.* at *4. The court also cited *Atkins v. State*, No. 05-07-0086-CR, 2008 WL 2815087 (Tex. Ct. App. July 23, 2008), which imposed a heightened penalty on an HIV-positive individual, despite the lack of evidence that the victim was exposed to HIV as a result of sexual assault.

82 Jody B. Gabel, *Liability for Knowing Transmission of HIV: The Evolution of a Duty to Disclose*, 21 FLA. ST. U. L. REV. 981, 1004 (1994).

83 *State v. Smith*, 621 A.2d 493 (N.J. Super. Ct. App. Div. 1993).

he knew that HIV could not spread through biting and spitting. He contended that he had come to this conclusion through HIV counseling sessions with a health care professional. He claimed that he threatened to kill others with HIV because he simply wanted to take advantage of others' ignorant beliefs. However, the appellate court rejected these arguments and held that his threatening speech amounted to the requisite intent.⁸⁴ The court accordingly affirmed the superior court's decision.⁸⁵

The fact that HIV can rarely be transmitted through biting and spitting⁸⁶ has not been found to negate the criminal intent of HIV-positive individuals. In *Weeks v. State*, an inmate spit twice in the face of an officer and was charged with attempted murder.⁸⁷ In this case, expert witnesses testified that the likelihood of transmitting HIV through spitting is extremely low.⁸⁸ The jury interpreted the expert testimony to mean that HIV could still be transmitted through spitting and convicted the defendant.

HIV-positive individuals are also convicted of aggravated assault after biting or spitting on others.⁸⁹ Courts have held that HIV is a deadly weapon whether or not the victims contract HIV from the conduct of HIV-positive defendants. In *Degrade v. State*, for example, an HIV-positive inmate bit an officer, who continued to test HIV-negative after the incident.⁹⁰ The court held that the mouth of an HIV-positive individual is a "deadly weapon,"⁹¹ which can "potentially" transmit HIV to the person being bitten and affirmed the trial court's conviction of aggravated assault.⁹²

E. Arguments Surrounding Criminalization of HIV Exposure and Available Evidence

Two opposite points have been raised with regards to the public health effect of criminalizing behavior that exposes others. Some argue that criminal sanctions promote public health by deterring risky sexual activities that spread HIV. Others believe that criminalization undermines HIV prevention because it may discourage HIV testing. This section explains both sides of this argument and introduces empirical evidence in support of each.

84 *Id.* at 509-10.

85 *Id.* at 493.

86 See *HIV & the Criminal Law: Low Risks, Theoretical Risks, and Impossible Routes*, NAM AIDSMAP, <http://www.aidsmap.com/Low-risks-theoretical-risks-and-impossible-routes/page/1320695> (last visited Dec. 18, 2013).

87 *Weeks v. State*, 834 S.W.2d 559 (Tex. Ct. App. 1992).

88 For example, one expert witness "stated that anything is theoretically possible but that the chance of transmitting HIV through saliva is the lowest in theoretical possibility." *Id.* at 564.

89 For the definition of aggravated assault, see *supra* note 77.

90 *Degrade v. State*, No. 05-04-00218-CR, 2005 Tex. App. LEXIS 547 (Jan. 26, 2005).

91 *Id.* at *5-6.

92 *Id.* at *3.

1. Arguments Advocating Criminalization of Behavior that Exposes Others to HIV

i. Deterrence Argument Supporting Criminalization

Sexual activities and needle sharing among injecting drug users are the most common causes of new HIV infections in the United States.⁹³ Since there is no vaccine or cure for HIV at present, the most effective way to stem the spread of HIV is to reduce the number of HIV-positive individuals who engage in these risky activities.⁹⁴

Supporters of criminalization claim that criminal punishment of risky activities deters at-risk individuals from engaging in them. The theory of deterrence in criminology generally assumes that a would-be offender will make a rational decision by comparing the benefits and costs of committing a crime.⁹⁵ According to supporters, HIV-positive individuals are more likely to avoid exposing others to HIV if they know that HIV-exposing behavior can result in criminal liability.⁹⁶

In fact, the 1988 report of the Presidential Commission on the Human Immunodeficiency Virus Epidemic recommended criminalization of HIV-exposing behavior, stating that “[e]stablishing criminal penalties for failure to comply with clearly set standards of conduct can also deter HIV-infected individuals from engaging in high-risk behaviors, thus protecting society against the spread of the disease.”⁹⁷ Based on this expectation, state legislatures adopted HIV-specific criminal statutes prohibiting HIV-exposing behavior. For example, California introduced criminal law provisions in 1998 to prohibit behavior that exposes others to HIV.⁹⁸ When the Senate Committee on Public Safety discussed the potential impact the bill might have on public health, proponents of the bill argued that the criminalization would deter high-risk behavior, thereby stopping the spread of HIV.⁹⁹ Despite civic groups’ counterpoint that the law might

93 *HIV/AIDS: Risk Behavior*, CTRES. FOR DISEASE CONTROL & PREVENTION (CDC), <http://www.cdc.gov/hiv/risk/behavior/index.html> (last updated May 31, 2013).

94 *Id.*

95 Richard Elliot, *Criminal Law, Public Health, and HIV Transmission: A Policy Options Paper*, UNAIDS 21 (June 2002), http://data.unaids.org/publications/IRC-pub02/jc733-criminallaw_en.pdf.

96 David Kromm, *HIV-Specific Knowing Transmission Statutes: A Proposal to Help Fight an Epidemic*, 14 ST. JOHN’S J. LEGAL COMMENT 253 (1999); Eric L. Schulman, *Sleeping with the Enemy: Combating the Sexual Spread of HIV-AIDS Through a Heightened Legal Duty*, 29 JOHN MARSHALL L. REV. 957, 977 (1996).

97 See PRESIDENTIAL COMMISSION REPORT, *supra* note 13, at 152.

98 CAL. HEALTH & SAFETY CODE §§ 120290-120291 (West 2013).

99 SEN. COMM. ON PUB. SAFETY, BILL ANALYSIS, S.B. 705 (Cal. Apr. 22, 1997). Groups against criminalization submitted reports to the Committee arguing that the state should focus its efforts on education and prevention of HIV. The opponents stressed that the criminal law might infringe upon the privacy rights of HIV-positive individuals and be selectively applied to disadvantaged people who are already facing strong stigma.

discourage people from undergoing testing, the Committee adopted the law.¹⁰⁰

Many courts have also repeatedly argued that imposing criminal liability for HIV-exposure would stop the spread of HIV.¹⁰¹ In *State v. Whitfield*, the court held that the criminalization of HIV-exposing behavior bears a reasonable relationship to a legitimate state objective, which is to stop transmission of a “deadly disease.”¹⁰² In some cases, courts have phrased the state’s interest in criminalizing HIV-exposing behavior as the protection of human life against the threat of HIV.¹⁰³

ii. Empirical Evidence Demonstrating the Failure of Criminal Law to Modify Risky Sexual Behavior of At-Risk Individuals

Contrary to the expectations of commentators, courts, and legislatures, public health experts have been doubtful about the deterrent effect of criminalization. These experts point out that the fear of criminal punishment is not likely to affect sexual decisions because sexuality is highly complex and involves many different feelings and desires.¹⁰⁴ Public health experts note that historically, laws regulating sexual behavior, such as sodomy statutes, have not effectively deterred such behavior.¹⁰⁵

Researchers have conducted empirical studies in order to examine the validity of each side of this argument. Three studies have found that criminalization fails to effectively deter the risky sexual activities of at-risk individuals. Scott Burris and colleagues conducted an empirical study comparing the behavior of at-risk individuals in a state with an HIV-specific criminal statute to behavior in a state without such a statute.¹⁰⁶ The researchers found that the sexual behavior of individuals in these two states differed little. Similarly, Horvath’s recent study quantitatively assessing the impact of HIV-specific statutes on sexual behavior of at-risk individuals found that criminalization had little effect on sexual activity without protection and/or disclosure of HIV-

¹⁰⁰ *Id.*

¹⁰¹ *State v. Whitfield*, 134 P.3d 1203, 1212 (Wash. App. Div. 2006).

¹⁰² *Id.* Many other states have ruled in favor of criminalizing behavior exposing others to HIV based on the same reason. *See, e.g.*, *State v. Musser*, 721 N.W.2d 734 (Iowa 2006); *State v. Gamberella*, 633 So. 2d 595 (La. Ct. App. 1993).

¹⁰³ In reviewing the constitutionality of an HIV-specific statute, the court stated that “essential to the analysis in the case at bar is the Supreme Court’s recognition and affirmation of the state’s compelling interest in protecting life. . . . Furthermore, a State [has] an unqualified interest in the preservation of human life.” *People v. Jensen*, 231 Mich. App. 439, 456 (Mich. Ct. App. 1998) (quoting *Roe v. Wade*, 410 U.S. 113, 153 (1973) (internal quotation marks omitted)).

¹⁰⁴ GOSTIN, *supra* note 70, at 189; Lawrence O. Gostin & James G. Hodge, Jr., *Piercing the Veil of Secrecy in HIV/AIDS and Other Sexually Transmitted Diseases: Theories of Privacy and Disclosure in Partner Notification*, 5 DUKE J. GENDER L. & POL’Y 9 (1998); Wolf & Vezina, *supra* note 19, at 873; Elliot, *supra* note 95, at 6.

¹⁰⁵ RICHARD D. MOHR, *GAYS/ JUSTICE: A STUDY OF ETHICS, SOCIETY, AND LAW* 51-53 (1988).

¹⁰⁶ Burris et al., *supra* note 7, at 467-68.

positive status.¹⁰⁷ The results show that residents of different states with different HIV criminal laws did not hold different attitudes towards HIV-positive individuals who have risky intercourse without disclosure.¹⁰⁸ In a qualitative study, researchers conducted interviews with homosexual men with HIV in the United Kingdom to investigate their experiences of unprotected anal intercourse.¹⁰⁹ The findings of this study suggest that the threat of criminal prosecution drove some people to disclose HIV status or take precautions, but the law also moved others towards “increased anonymity” in sexual relationships and reduced openness about HIV status, which could be detrimental to HIV prevention.¹¹⁰

2. Criticisms Against Criminalization as a Disincentive for HIV Testing

i. Public Health Concerns Involving HIV Testing

Public health experts and advocacy groups are worried that criminalization of HIV-exposing behavior might discourage individuals at high risk of HIV from testing.¹¹¹ Although there is an extremely small portion of HIV-positive individuals who use their virus to harm others, accumulated evidence shows that most HIV-positive individuals tend to protect others once they learn that they are HIV-positive.¹¹² Citing this evidence, public health experts stress the importance of at-risk individuals learning their HIV status through testing at the earliest

¹⁰⁷ Horvath et al., *supra* note 7, at 1225.

¹⁰⁸ The researchers used the participant’s residency in a state with an HIV-specific statute as a proxy for whether criminalizing law governed the participant’s behavior. Trained interviewers interviewed survey participants in a 3.5-month period in 2008 using online banner advertisements placed on two websites popular with gay and bisexual men. The banner ad stated, “Participate in University Research on Sex and Alcohol and Earn \$30.” It included the university and study logo and a picture of a man. *Id.* at 1222. A total of 1725 participants, who identified themselves as male and having had sex with a man, completed the survey.

¹⁰⁹ Dodds et al., *supra* note 7, at 137.

¹¹⁰ *Id.* at 142. Of 29 men who reflected on personal impact, almost half felt that prosecutions had not influenced their sexual behavior in any way. The rest said they had planned to behave and communicate differently with their sex partners as a direct result of concern at “the prospect of legal intrusion into their sex lives.” *Id.* at 140.

¹¹¹ *Criminalization of HIV Transmission*, UNAIDS 4-5 (Aug. 2008), http://data.unaids.org/pub/basedocument/2008/20080731_jc1513_policy_criminalization_en.pdf; *Verdict on a Virus: Public Health, Human Rights and Criminal Law*, INT’L PLANNED PARENTHOOD FEDERATION (IPPF) 25 (Nov. 13 2008), http://www.ippf.org/sites/default/files/verdict_on_a_virus.pdf.

¹¹² Nicole Crepaz et al., *Do Prevention Interventions Reduce HIV Risk Behaviours Among People Living with HIV? A Meta-Analytic Review of Controlled Trials*, 20 AIDS 143, 144 (2006); Peter H. Kilmarx, Francoise F. Hamers, & Thomas A. Peterman, *Living with HIV: Experiences and Perspectives of HIV-Infected Sexually Transmitted Disease Clinic Patients After Posttest Counseling*, 15 SEXUALLY TRANSMITTED DISEASES 25 (1998); Gary Marks, Nicole Crepaz & Robert S. Janssen, *Estimating Sexual Transmission of HIV from Persons Aware and Unaware That They Are Infected with the Virus in the USA*, 20 AIDS 1447 (2006); Weinhardt et al., *supra* note 6, at 1403.

possible time.¹¹³

Currently, in the United States, it is estimated that approximately one in five people infected with HIV (21%) are unaware that they are infected and may be “unknowingly transmitting the virus to others.”¹¹⁴ HIV has an asymptomatic incubation period that can last up to 10 years after infection, leaving HIV-positive individuals unaware of their HIV-positive status unless they are tested.¹¹⁵ HIV-positive individuals are infectious during this asymptomatic period, especially if they do not receive antiretroviral treatments.¹¹⁶ Public health efforts have responded to this situation by focusing on increasing HIV testing of at-risk individuals to inform them of their status at the earliest possible time.¹¹⁷

Public health experts worry that the criminalization of HIV-exposing behavior undermines these efforts because at-risk individuals may not undergo testing due to fear of criminal charges. The majority of HIV-specific criminal statutes in the United States punish HIV-positive individuals for engaging in certain behaviors if they know that they are HIV-positive.¹¹⁸ According to public health experts, individuals at high risk of infection may avoid testing because awareness of their status could be used against them in court.¹¹⁹

ii. Media Amplification of Criminalization's Negative Effect on HIV Testing

According to existing research, media coverage can increase the impact of criminal law on HIV testing in several ways. First, media coverage can raise public awareness of the law. Because many people do not have direct experience with the criminal justice system, the public's knowledge of the law may rely

113 Gary Marks, Scott Burris & Thomas A. Peterman, *Reducing Sexual Transmission of HIV from Those Who Know They Are Infected: The Need for Personal and Collective Responsibility*, 13 AIDS 297 (1999).

114 *HIV/AIDS: Risk Behavior*, *supra* note 93, at 2.

115 *HIV/AIDS: HIV Basics*, *supra* note 56.

116 T. Deidre Hollingsworth et al., *HIV-1 Transmission, by Stage of Infection*, 198 J. INFECTIOUS DISEASE 687 (2008); Christopher D. Pilcher et al., *Brief but Efficient: Acute HIV Infection and the Sexual Transmission of HIV*, 189 J. INFECTIOUS DISEASE 1785 (2004); Maria J. Wawer et al., *Rates of HIV-1 Transmission Per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda*, 191 J. INFECTIOUS DISEASE 1403 (2005).

117 Granich et al., *supra* note 6; *WHO and UNAIDS Issue New Guidance on HIV Testing and Counselling in Health Facilities*, WORLD HEALTH ORG. (WHO) (2007), <http://www.who.int/mediacentre/news/releases/2007/pr24/en/index.html>; *Act Against AIDS: Testing*, CDC, <http://www.cdc.gov/actagainstaids/basics/testing.html> (last updated Nov. 20, 2013); *UNAIDS/WHO Policy Statement on HIV Testing*, UNAIDS & WHO (June 2004), http://www.who.int/rpc/research_ethics/hivtestingpolicy_en_pdf.pdf.

118 *See supra* Section I.B.

119 Donald H.J. Hermann, *Criminalizing Conduct Related to HIV Transmission*, 9 ST. LOUIS U. PUB. L. REV. 351, 356-57 (1990); Kenney, *supra* note 17, at 272; Robert Klitzman, et al., *Naming Names: Perceptions of Name-Based HIV Reporting, Partner Notification, and Criminalization of Non-Disclosure Among Persons Living with HIV*, SEXUALITY RES. & SOC. POL'Y, Sept. 2004, at 38, 49-50 (2004); *Verdict on a Virus*, *supra* note 111, at 37.

largely on how the law is presented in the media.¹²⁰ When the media publicizes that criminal courts use HIV test results to prove that HIV-positive individuals knowingly exposed others, at-risk individuals may decide to avoid testing to stay “legally negative.”¹²¹ An empirical study investigated how at-risk individuals become aware of criminal liability for exposing others. Respondents pointed to media coverage as one of few information sources.¹²²

Second, the media could also indirectly discourage at-risk individuals’ testing by cultivating a negative impression of HIV-positive individuals. Public health advocates claim that the media tends to highlight negative, sensational aspects of criminal cases, such as an HIV-positive criminal having malicious intention to infect others.¹²³ According to these advocates, the sensationalized coverage could reinforce negative attitudes that society may already have against the HIV-positive population.¹²⁴ Public health experts have pointed out that such social hostility could undermine public health efforts to encourage voluntary HIV testing and counseling of at-risk individuals.¹²⁵

iii. Dearth of Empirical Evidence to Assess Criminalization’s Impact on HIV Testing Decisions

The validity of concerns regarding criminalization has not been empirically tested. To fill this research gap, this Note conducts a quantitative study to examine the impact of criminal laws on HIV testing decisions. The results of regression analysis fail to support the claim that people are less likely to be tested for HIV when they are subject to HIV-specific criminal law that prohibits exposure of HIV. The regression analysis, however, supports the claim that

120 Valerie P. Hans & Juliet L. Dee, *Media Coverage of Law: Its Impact on Juries and the Public*, 35 AM. BEHAV. SCIENTIST (1991); Jennifer K. Robbennolt & Christina A. Studebaker, *News Media Reporting on Civil Litigation and Its Influence on Civil Justice Decision Making*, 27 L. & HUM. BEHAV. 5 (2003).

121 Burris et al., *supra* note 7, at 512.

122 Carol L. Galletly et al., *HIV-Positive Persons’ Awareness and Understanding of Their State’s Criminal HIV Disclosure Law*, 13 AIDS BEHAV. 1262 (2009).

123 Ralf Jürgens et al., *Ten Reasons to Oppose the Criminalization of HIV Exposure or Transmission*, 17 REPROD. HEALTH MATTERS 163, 166 (2009); *Verdict on a Virus*, *supra* note 111 at 24-26; *Criminal Law, Public Health, and HIV Transmission*, *supra* note 95, at 23-24. For a collection of media reports on criminalization of behavior exposing others to HIV, see CRIM. HIV TRANSMISSION, <http://criminalhivtransmission.blogspot.kr/> (last visited Dec. 15, 2013); *What Impact Does the Media Have?*, IPPF, <http://www.hivandthelaw.com/basic-information/ten-questions/what-impact-does-media-have> (last visited Dec. 15, 2013).

124 Jürgens et al., *supra* note 123, at 166; Elliot, *supra* note 95, at 23-24; *Verdict on a Virus*, *supra* note 111, at 24-26.

125 Ron Stall et al., *Decisions to Get HIV Tested and to Accept Antiretroviral Therapies Among Gay/Bisexual Men: Implications for Secondary Prevention Efforts*, 11 J. AIDS 151 (1996); Ronald O. Valdiserri, *HIV/AIDS Stigma: An Impediment to Public Health*, 92 AM. J. PUB. HEALTH 341, 341-42 (2002); Ronald O. Valdiserri et al., *Promoting Early HIV Diagnosis and Entry into Care*, 13 AIDS 2317, 2317-30 (1999); see F.M. Boyd et al., *What Do Pregnant Women Think About the HIV Test? A Qualitative Study*, 11 AIDS CARE 21 (1999).

increased reporting of criminalization of HIV exposure is associated with reduced HIV testing of at-risk individuals.

II. EMPIRICAL ANALYSIS OF HIV-SPECIFIC CRIMINAL STATUTES' IMPACT ON HIV TESTING BEHAVIOR OF AT-RISK INDIVIDUALS

A. Methodologies

1. Regression Analysis

The analysis is designed to test the null hypotheses that there would be no difference in self-reported HIV testing decisions in the past 12 months (1) between at-risk individuals living in states with HIV-specific statutes and those in states without a statute, and (2) between at-risk individuals living in states with HIV-specific statutes and intense media coverage of the law and at-risk individuals residing in other states.

The study utilizes the concept of “difference-in-differences” (DID) estimators using two groups of states: (1) states that introduced HIV-specific statutes between 2002 and 2009 (treatment states), and (2) states that did not adopt a new HIV-specific law during this period (control states).¹²⁶ The regression model takes the following form:

$$\text{Outcome}_{sti} = F(\beta_0 + \beta_1 * C_{st} + \beta_2 * X_{st} + \beta_3 * Y_{sti} + \psi_s + \varphi_t + \varepsilon_{st}),$$

where **Outcome_{sti}** is whether respondent *i* living in state *s* and interviewed at time *t* reports having taken HIV testing in the past 12 months, **C_{st}** is whether the state the survey participant was residing in at the time of interview had an HIV-specific criminal law, **X_{st}** is the HIV prevention policy and other environmental factors in the state the respondent was living in in the year of interview, **Y_{sti}** is a vector of the demographic, social, and economic characteristics of the survey participant, **ψ_s** is a set of state fixed-effect dummies, **φ_t** is a set of year fixed-effect dummies, and **F** is the logistic function.

i. Outcome Variable

The outcome variable is a binary indicator of whether the survey participant reported that he or she had been tested for HIV in the past 12 months. Due to the

¹²⁶ DID is an econometric models to analyze panel data that were collected before and after a treatment in a treatment group and a control group. DID estimators represent the average change in *Y* in the treatment group over the course of the experiment, minus the average change in *Y* in the control group over the same period. The merit of this estimation, compared to standard regression analysis, is that, by comparing the changes in two different groups, pretreatment differences in *Y* are eliminated. JAMES H. STOCK & MARK W. WATSON, INTRODUCTION TO ECONOMETRICS 480-82 (2d ed. 2007).

nature of the BRFSS, this study uses self-reported HIV testing experiences of survey participants. Because the dependent variable is binary, this study uses a logistic regression model.¹²⁷

ii. Treatment Variables Related to Criminalization of HIV-Exposing Behavior in Residing State

Table 3. Descriptions of Treatment Variables

Variable name	Descriptions
HIV-Specific Law	=1 if the state of residence had an HIV-specific criminal statute at the time of interview, 0 otherwise
HIV-Specific Law *Frequency of Media Reportage	Interaction term between the HIV-specific law variable and the frequency of news reports of HIV criminalization in that state in that year (continuous variable)

Dummy Variable to Indicate Whether the Survey Participant's State of Residence had an HIV-Specific Criminal Statute in the Year of Interview

This study investigates whether residing in a state that criminalizes HIV-exposing behavior through an HIV-specific statute affects HIV testing decisions. The judiciary has also applied traditional criminal law and STD law to HIV-exposing behavior, and the impacts of these laws are also worth investigating.¹²⁸ However, there is no reliable data on the number of prosecutions of HIV-exposing behavior based on traditional criminal law and STD law in each state by year. This study, therefore, focuses on the impact of HIV-specific criminal statutes.

The HIV-specific law variable takes the value of 1 if the survey participant's state of residence had an HIV-specific criminal law in the year of interview. No state has abolished HIV-specific laws, so the HIV-specific law variable takes the value of 1 for all years from the year the law was adopted.¹²⁹ For the states that did not adopt an HIV-specific statute, this variable is coded 0.¹³⁰

¹²⁷ Logistic regression is a non-linear regression model that is specifically designed for binary dependent variables. For details, see *id.* at 389.

¹²⁸ See *supra* Sections I.C and I.D.

¹²⁹ The states that introduced an HIV-specific criminal law between 2002 and 2009 are Alaska and Mississippi.

¹³⁰ The states that did not adopt an HIV-specific statute are: Arizona, Connecticut, District of Columbia, Hawaii, Kansas, Maine, Massachusetts, Minnesota, Nebraska, New Hampshire, New Mexico, New York, Oregon, Rhode Island, Texas, Vermont, West Virginia, and Wyoming.

*Interaction Term Between the HIV-Specific Law Variable and the
Frequency of Media Reporting of Criminalization of HIV-Exposing
Behavior in the State of Residence in the Year of Interview Variable*

The regression model includes an interaction term between the HIV-specific law variable and the continuous variable that indicates the frequency of media reportage on criminalization of HIV-exposing behavior. The model includes this variable, based on previous research findings that it could be media reportage, rather than the actual law or its enforcement, which decides the influence of the law in society.¹³¹

This study focuses on printed media because newspapers and journals are an “accessible, non-transient form of media.”¹³² This study uses two databases to search for newspaper and journal articles that discuss criminalization of behavior that exposes others to HIV. For the newspaper and journal articles that the general public has access to, the Westlaw United States Newspaper (USNP) database is searched with several different combinations of keywords.¹³³ The USNP database contains news reports from United States Papers, as provided by NewsRoom to West, a Thomson Reuters business.

The second database used is the Ethnic NewsWatch (ENW) database,¹³⁴ which focuses on news sources popular with populations known to be at higher risk of HIV infection.¹³⁵ The ENW features newspapers and magazines from ethnic and minority presses. It presents a comprehensive, full-text collection of more than 2.5 million articles from more than 340 publications offering both national and regional coverage.¹³⁶ This database also includes reports presented in Spanish, such as *El Nuevo Herald*.

For this study, several different combinations of keywords were used to

¹³¹ See *supra* Subsection I.E.2.

¹³² Richard C. Adelman & Lois M. Verbrugge, *Death Makes News: The Social Impact of Disease on Newspaper Coverage*, 41 J. HEALTH & SOC. BEHAV. 347, 347 (2000).

¹³³ The combinations of keywords used are: “HIV” & “sentence!”, “HIV” & “crim!”, “HIV” & “accus!”, “HIV” & “prosecut!”, “AIDS” & “transmit!” & “crim!”, “AIDS” & “spread” & “crim!”, and “AIDS” & “expos!” & “crim!”.

¹³⁴ *Ethnic NewsWatch*, PROQUEST, http://www.proquest.com/en-US/catalogs/databases/detail/ethnic_newswatch.shtml (last visited Dec. 15, 2013).

¹³⁵ Because this study focuses on the behavior of individuals at elevated risk of HIV infection, this study also takes into consideration the media to which the high-risk population is most exposed. The CDC has repeatedly reported that individuals with non-white racial backgrounds constitute the majority of the HIV-positive population in the United States. Div. of HIV/AIDS Prevention, Nat’l Ctr. for HIV/AIDS, Viral Hepatitis, STD & TB Prevention, *HIV in the United States: At a Glance*, CDC 2 (Nov. 2013), http://www.cdc.gov/hiv/pdf/Statistics_Basics_Factsheet.pdf.

¹³⁶ *Ethnic NewsWatch*, *supra* note 134.

collect reports about criminalization of HIV-exposing behavior from the ENW.¹³⁷ This study counted the number of newspaper articles that discuss (1) criminal charges brought against HIV-exposing individuals or (2) debates surrounding an HIV-specific statute in that state. This study does not include articles about HIV-specific statutes in other states or countries.

In counting the frequency of reporting, articles published in national press outlets such as the *Chicago Tribune* and the *Los Angeles Times* are considered to have an impact across states in the year of reporting. Articles featured in regional or local press outlets are considered to have an influence limited to the state where the article was published.

Because this media reportage variable is created based on news reports available on Westlaw, this variable does not capture the frequency of reports in other media such as television and radio. Therefore, it might have limitations in serving as an accurate proxy for media coverage. But this study is valuable in that it provides at least a rough estimate and is the first study to attempt to empirically assess the effect of media on the law's impact.

iii. Variables Related to State-Level HIV Policies and Other Environmental Factors

Table 4. Descriptions of control variables

Variable name	Descriptions
CDC prevention funding	Preceding three-year average of CDC spending on HIV prevention in the state of residence
ADAPs spending	ADAPs budget per HIV patient in the state of residence in the interview year
AIDS rate	Preceding three-year average of annual AIDS rate per 100,000 population in the state of residence

Average of Preceding Three Years' CDC Spending on HIV Prevention Programs in the Respondent's State of Residence

The regression model includes a continuous variable to reflect the CDC's spending on HIV prevention programs in the respondent's state of residence around the time of interview. This CDC spending variable serves as a proxy for various factors related to states' HIV prevention programs that affect HIV testing behavior. Studies indicate that the way local HIV prevention programs are

¹³⁷ The combinations of keywords "HIV" & "crim!", "HIV" & "prosecut!", and "HIV" & "accus!" were used to search for articles that featured criminalization of behavior involving exposing others to HIV.

developed and run affect HIV testing behavior. For example, HIV testing decisions are influenced by access to HIV testing, the variety of test types offered, whether someone such as a health care provider or local HIV groups actively encouraged HIV testing, and whether an individual was exposed to information that increased understanding of how one can acquire HIV and prevent or treat infection.¹³⁸

However, comprehensive data about local HIV testing programs are unavailable. The BRFSS surveys conducted in recent years contain information about some of the issues at stake, such as what type of HIV testing was offered and whether the testing site was easily accessible, but the BRFSS data collected in earlier years do not have information about these factors. This study, therefore, uses the CDC's funding for state and local HIV prevention programs as a proxy of how actively the prevention programs were run, how accessible HIV testing was, whether a testing clinic had various options that suited the needs and preferences of individuals being tested, and whether the state had active education programs to encourage HIV testing of at-risk individuals.

The CDC's funding for state and local HIV prevention programs is chosen as a proxy of state and local HIV prevention programs because the CDC's funding significantly affects state and local HIV testing programs. The CDC fund constitutes a large portion of state governments' spending on HIV prevention activities. Unpublished reports from 40 states to the CDC in 2000 indicate that federal funding accounted for approximately 60% of total HIV prevention spending in those states, ranging from 25% to 100% in each state.¹³⁹ The Kaiser Family Foundation (KFF) estimates that CDC funding constituted between 84% and 95% of federal funding between the fiscal years 1995 and 2004.¹⁴⁰ Reflecting

138 See Joseph N. Inungu, *Potential Barriers to Seeking Human Immunodeficiency Virus Testing Among Adults in the United States: Data from the 1998 National Health Survey*, 16 AIDS PATIENT CARE & STDs 293 (2002); Scott E. Kellerman et al., *HIV Testing Within At-Risk Populations in the United States and the Reasons for Seeking or Avoiding HIV Testing*, 31 J. AIDS 202 (2002); Freya Spielberg et al., *Choosing HIV Counseling and Testing Strategies for Outreach Settings, a Randomized Trial*, 38 AIDS 348 (2005) [hereinafter Spielberg et al., *Testing Strategies*]; Freya Spielberg et al., *Moving from Apprehension to Action: HIV Counseling and Testing Preferences in Three At-Risk Populations*, 13 AIDS EDUC. & PREVENTION 524 (2001) [hereinafter Spielberg et al., *Testing Preferences*]; Freya Spielberg et al., *Overcoming Barriers to HIV Testing: Preferences for New Strategies Among Clients of a Needle Exchange, a Sexually Transmitted Disease Clinic, and Sex Venues for Men Who Have Sex with Men*, 32 J. AIDS 318 (2003) [hereinafter Spielberg et al., *Overcoming Barriers*]; Catherine Worthington & Ted Myers, *Desired Elements of HIV Testing Services: Test Recipient Perspectives*, 16 AIDS PATIENT CARE & STDs 537 (2002).

139 Harrel W. Chesson et al., *Does Funding for HIV and Sexually Transmitted Disease Prevention Matter? Evidence from Panel Data*, 29 EVALUATION REV. 3 (2005).

140 Todd Summers & Jennifer Kates, *Trends in U.S. Government Funding for HIV/AIDS—Fiscal Years 1981 to 2004*, KAISER FAMILY FOUND. 8 (2004), <http://kaiserfamilyfoundation.files.wordpress.com/2013/01/issue-brief-trends-in-US-government-funding-for-HIV-aids-fiscal-years-1981-to-2004.pdf>; see also David R. Holtgrave & Jennifer Kates, *HIV Incidence and CDC's HIV Prevention Budget An Exploratory Correlational Analysis*, 32 AM. J. PREVENTIVE MED. 63, 63-64 (2007) (describing personal communication with a CDC

the significance of CDC funding for HIV testing programs, a team of researchers reported that changes in the CDC's funding of state HIV prevention programs correlated with changes in the rate of people being tested for HIV.¹⁴¹

In creating this CDC spending variable, this study refers to the methodology used by Linas et al. (2006).¹⁴² Data of the National Alliance of State and Territorial AIDS Directors (NASTAD), which tracked the annual CDC funding for HIV prevention provided to state and local governments, is adjusted by the Consumer Price Index (CPI) for Urban Consumers in four different regions (Northeast, West, Midwest, and South).¹⁴³ Subsequently, the average spending of the past three years, including the year of the interview, is entered into the regression model as a continuous variable.

ADAPs' Budget Per HIV Patient in the Respondent's State of Residence in the Interview Year

The regression includes a variable, which represents spending by AIDS Drug Assistance Programs (ADAPs) for one HIV patient in the respondent's state of residence in the interview year. This continuous variable is a proxy for the availability of HIV treatment in the event an individual tests HIV-positive.

Many studies have found that people are more likely to be tested for HIV if treatments are available.¹⁴⁴ Highly Active Antiretroviral Treatment (HAART) has been used to treat HIV-positive patients since the mid-1990s and has led to a dramatic prolongation of healthy lives of HIV-positive individuals.¹⁴⁵ HAART, however, has not been available to everyone in need of treatment due to its high cost.¹⁴⁶

The ADAPs have provided HIV-related prescription drugs to low-income HIV patients who would otherwise have limited or no coverage for prescription

officer indicating that CDC funds constituted 84% of federal HIV/AIDS funds for prevention in the fiscal year 2005).

141 Benjamin P. Linas et al., *Assessing the Impact of Federal HIV Prevention Spending on HIV Testing and Awareness*, 96 AM. J. PUB. HEALTH 1038, 1040 (2006).

142 *Id.* at 1038-39.

143 NASTAD data was provided by NASTAD and is on file with the author.

144 Kellerman et al., *supra* note 138; Spielberg, et al., *Overcoming Barriers*, *supra* note 138; Spielberg, et al., *Testing Preferences*, *supra* note 138, at 524.

145 *Evolution of HIV/AIDS Prevention Programs—United States, 1981-2006*, 55 MORBIDITY & MORTALITY WKLY. REP. 597 (2006); Pythia T. Nieuwkerk et al., *Quality of Life in Asymptomatic- and Symptomatic HIV Infected Patients in a Trial of Ritonavir/Saquinavir Therapy*, 14 J. AIDS 181 (2000); CASCADE Collaboration, *Determinants of Survival Following HIV-1 Seroconversion After the Introduction of HAART*, 362 LANCET 1267 (2003).

146 *Selected Measures of Access to and Utilization of Treatment and Prophylaxis for HIV-Infected Persons*, 9 HIV/AIDS SURVEILLANCE SUPPLEMENTAL REP. 1 (2003), available at http://www.cdc.gov/hiv/pdf/statistics_2003_HIV_Surveillance_Report_vol_9_no3.pdf.

drugs.¹⁴⁷ The ADAPs began in 1987 and grew in the number of clients they serve. With more than 134,000 enrollees as of 2006, the ADAPs reach approximately one-quarter of those individuals with HIV estimated to be receiving care in the United States. In June of 2005 alone, the ADAPs provided medications to more than 96,000 clients and insurance coverage to thousands more. The ADAPs operate in all 50 states and the District of Columbia. The Ryan White Care Act requires all of ADAPs' clients to be HIV-positive as well as low income and either uninsured or underinsured.¹⁴⁸ Reflecting the purpose of the programs, for most years, the majority of ADAPs' beneficiaries have been people of color and the uninsured.¹⁴⁹

The regression model includes a variable for ADAPs' budget per HIV patient in the respondent's state of residence in the year of interview. This variable is expected to function as a proxy for how likely treatment would be available for people who test positive for HIV. This study first collects each state's ADAPs spending from 2002 to 2009 from the KFF reports and then adjusts it by the regional CPI for Urban Consumers to reflect inflation along the years in four different regions (Northeast, South, West, and Midwest).¹⁵⁰ This adjusted amount is then divided by the average AIDS cases in that state in three preceding years. The three preceding years' average AIDS cases are a proxy of the prevalence of HIV in that state. By dividing the amount of ADAPs spending by the number of average AIDS cases, this variable indicates the average money spent for each HIV patient in that state in that year. This variable is included in the regression model as a continuous variable.

Average of Preceding Three Years' AIDS Rates Per 100,000 People in the Respondent's State of Residence

Previous findings suggest that the number of people tested for HIV is associated with the prevalence of HIV in the community; the higher the prevalence, the more likely people are to be tested.¹⁵¹ Because there is no reliable

147 The description of ADAPs can be found at *AIDS Drug Assistance Programs (ADAPs)*, KAISER FAMILY FOUND. (June 25 2013), <http://kff.org/hiv/aids/fact-sheet/aids-drug-assistance-programs/>.

148 42 U.S.C.A. § 300ff-1 (2006); see also Nat'l Alliance of State & Territorial AIDS Dirs. et al., *National ADAP Monitoring Project Annual Report: Module 1*, NAT'L ALLIANCE STATE & TERRITORIAL AIDS DIRECTORS (Jan. 2013), <http://www.nastad.org/Docs/NASTAD-National-ADAP-Monitoring-Project-Report-Module-1-2013-1.pdf>.

149 Nat'l Alliance of State & Territorial AIDS Dirs. et al., *supra* note 148.

150 *Id.*

151 Peter D. Ehrenkranz et al., *Written Informed-Consent Statutes and HIV Testing*, 37 AM. J. PREVENTIVE MED. 37, 60 (2009) (reporting that respondents living in states with a higher prevalence of AIDS per capita were more likely to report recent HIV testing (18.3%) than people living in states with a lower prevalence of AIDS per capita (12.5%)); Kathryn A. Phillips, *The Relationship of 1988 State HIV Testing Policies to Previous and Planned Voluntary Use of HIV Testing*, 7 J. AIDS 403, 405 (1994) (reporting that testing rates generally increase with higher AIDS incidence).

data on HIV prevalence in each state in each year, this study uses the average of three preceding years' AIDS rates per 100,000 people in the respondent's state of residence as a proxy of HIV prevalence. This variable is expected to control for any influence a high prevalence of HIV might have on an individual's decision to be tested.

This study uses publicly available CDC data to calculate the average of three year's AIDS rate per 100,000 people from 2002 to 2009.¹⁵² The regression includes this average AIDS rate in the state the survey participant lived in as a continuous variable.

iv. Variables Related to Individual Characteristics of Survey Participants

Table 5. Factors Included in Regression and Factors Reported as Potential Confounder

Factors Reported to Affect HIV Testing	Factors in Regression Model
Race ¹⁵³	Race
Age ¹⁵⁴	Age
Sex ¹⁵⁵	Sex
Marital status ¹⁵⁶	Marital status
Pregnancy experience ¹⁵⁷	[Proxies] Age, Sex, Martial status, Current pregnancy status
Education level ¹⁵⁸	Education level
Employment status ¹⁵⁹	Employment status
Income level ¹⁶⁰	Income level

¹⁵² *HIV/AIDS: HIV Surveillance Reports*, CDC, <http://www.cdc.gov/hiv/library/reports/surveillance/> (last updated Nov. 27, 2013).

¹⁵³ See Inungu, *supra* note 138; Paul A. Simon et al., *Reasons for HIV Antibody Test Refusal in a Heterosexual Sexually Transmitted Disease Clinic Population*, 10 AIDS 1549 (1996); Kellerman et al., *supra* note 138.

¹⁵⁴ See Olga A. Grinstead et al., *Antibody Testing and Condom Use Among Heterosexual African Americans at Risk for HIV Infection: The National AIDS Behavioral Surveys*, 87 AM. J. PUB. HEALTH 857 (1997); Inungu, *supra* note 138; Kellerman et al., *supra* note 138.

¹⁵⁵ Kellerman et al., *supra* note 138; Simon et al., *supra* note 153.

¹⁵⁶ Grinstead et al., *supra* note 154; Inungu, *supra* note 138.

¹⁵⁷ See Jeffrey H. Samet et al., *Factors Associated With HIV Testing Among Sexually Active Adolescents: A Massachusetts Survey*, 100 PEDIATRICS 371 (1997).

¹⁵⁸ *Id.*

¹⁵⁹ *Id.*

¹⁶⁰ Spielberg et al., *Testing Strategies*, *supra* note 138, at 349. Many studies have limited their scope of analysis to HIV testing of at-risk individuals with low-income. See Angela B. Hutchinson et al., *Understanding the Patient's Perspective on Rapid and Routine HIV Testing in an Inner City Urgent Care Center*, 16 AIDS EDUC. & PREVENTION 101, 112 (2004); Kathleen J. Sikkema et al., *Outcomes of a Randomized Community-Level HIV Prevention Intervention for Women Living in 18 Low-income Housing Developments*, 90 AM. J. PUB. HEALTH 57 (2000).

No time/money for HIV testing ¹⁶¹	[Proxies] Health insurance coverage, Employment status, Income level
Felt health problem related to HIV ¹⁶²	[Proxy] General health condition

The regression model includes the demographic and socioeconomic characteristics of survey participants as control variables. The variables include age, race, sex, marital status, current pregnancy status, education level, employment status, income level, health insurance coverage, and self-perception of general health condition. In selecting the variables, this study refers to other empirical studies that include a similar list of control variables to assess the impact of a public policy on HIV testing behavior.¹⁶³ Many studies have found that a number of demographic and socio-economic factors affect HIV testing behavior.¹⁶⁴

For factors not included in the BRFSS, this study uses proxies to control for potential confounders. The BRFSS does not have information about whether the survey participant had ever been pregnant, whether the participant had financial barriers to HIV testing, or if the participant experienced an HIV-related health condition. For pregnancy experience, this study uses age, sex, marital status, and current pregnancy status as proxies. For financial barriers to HIV testing, this study uses health insurance coverage, employment status, and income level as proxies. For whether the person felt that he or she had a health problem associated with an HIV infection, this study uses self-perceived general health condition as a proxy.

Table 6. Descriptions of Control Variables

Controlled Factor	Variables	Descriptions
Race ¹⁶⁵	Race: White	=1 if yes, 0 otherwise
	Race: Black	
	Race: Hispanic	=1 if value missing, 0 otherwise
	Race missing	
Age	Age ¹⁶⁶	Continuous variable
Sex	Sex	=1 if male, 0 if female
Now pregnant	Pregnancy	=1 if pregnant, 0 otherwise
	Pregnancy missing	=1 if value missing, 0 otherwise

161 Spielberg, et al., *Overcoming Barriers*, *supra* note 138; Spielberg et al. *Testing Preferences*, *supra* note 138.

162 Ted Myers et al., *Factors Affecting Gay and Bisexual Men’s Decisions and Intentions to Seek HIV Testing*, 83 AM. J. PUB. HEALTH 701 (1993); Kellerman, *supra* note 138.

163 Ehrenkranz, *supra* note 151; Linas, *supra* note 141; Phillips, *supra* note 151; Nicola M. Zetola et al., *Association Between Rates of HIV Testing and Elimination of Written Consents in San Francisco*, 297 JAMA 1061 (2007).

164 Inungu, *supra* note 138; Grinstead, *supra* note 154; Simon, *supra* note 153.

165 “Other races” includes “Asian,” “Pacific Islander,” “American Indian,” and “Alaska Native.”

166 Due to the design of the BRFSS survey, the samples consist of individuals aged 18 and older.

Marital status	Married ¹⁶⁷	=1 if currently married, 0 otherwise
	Married missing	=1 if value missing, 0 otherwise
Education level	Education level 1 _4 ¹⁶⁸	} =1 if yes, 0 otherwise
	Education level 2 _4 ¹⁶⁹	
	Education level 3 _4 ¹⁷⁰	
	Education level missing	=1 if value missing, 0 otherwise
<hr/>		
Employment status	Employed ¹⁷¹	=1 if yes, 0 otherwise
	Employed missing	=1 if value missing, 0 otherwise
Income level	Income level 1 _4 ¹⁷²	} =1 if yes, 0 otherwise
	Income level 2 _4 ¹⁷³	
	Income level 3 _4 ¹⁷⁴	
	Income level missing	=1 if value missing, 0 otherwise
Health insurance	Have health insurance	=1 if yes, 0 otherwise
	Health insurance missing	=1 if value missing, 0 otherwise
General health condition	Poor health ¹⁷⁵	=1 if yes, 0 otherwise
	Poor health missing	=1 if value missing, 0 otherwise

(*: Excluded from regression to avoid multicollinearity)

Each variable is created as presented in Table 6. In order to capture any bias created by missing values, this study includes a separate binary variable to indicate cases with missing values for each control variable. These variables are coded 1 if the value is missing and 0 if the question is answered in any way.

The regression model does not include a variable for the participant's self-perceived risk of HIV infection, although other studies have reported that this

167 The answers classified as "unmarried" are: "divorced," "widowed," "separated," "never been married," and "a member of an unmarried couple."

168 The answers that belong to this category are: "Never attended school or only kindergarten," "Grades 1-8 (Elementary)," and "Grades 9-11 (Some high school)."

169 This variable represents whether the survey participant reports having had "Grade 12 or GED (High school graduate)."

170 This variable represents whether the survey participant reports having had "education of College 1-3 years (Some college or technical school)."

171 The answers classified as employed are: "employed for wages" and "self-employed." The answers classified as unemployed are: "out of work for more than 1 year," "out of work for less than 1 year," "homemaker," "student," "retired," and "unable to work."

172 The answers classified under this category are: "under 10,000," "10,000 to less than 15,000," and "15,000 to less than 20,000."

173 The answers classified under this category are: "20,000 to less than 25,000" and "25,000 to less than 35,000."

174 The answers classified under this category are: "35,000 to less than 50,000" and "50,000 to less than 75,000."

175 The answers classified as non-poor health are: "excellent," "very good," "good," or "fair" health.

factor substantially influences HIV testing decisions.¹⁷⁶ The variable is not included because this study uses cases where the survey participant acknowledged being in a high-risk situation.¹⁷⁷ Consequently, there is no relevant difference in the perception of HIV infection risk among survey respondents included in the analysis.

Because the risk acknowledgement question asks about the respondent's experience in engaging in a risky sexual activity, the potential confounders related to sexual activities are controlled for in the regression analysis. They include the experience of having had more than one sex partner, intercourse with a non-regular partner, and unprotected vaginal or anal intercourse.¹⁷⁸

This regression does not include control variables about emotional factors because the BRFSS does not include information about these factors. According to research, an individual's decision to undertake HIV testing could be affected by fear of learning about HIV, dislike of needles, stress of waiting for the test result, lack of partner or peer support, lack of motivation to protect others, and dislike of condom use.¹⁷⁹ However, omission of these emotional factors in the regression analysis is not likely to undermine the reliability of the estimates of primary independent variables because these factors are not systematically associated with the HIV-specific law variable; there is no reason to believe that at-risk individuals' characteristics along these dimensions are different in states that have criminal exposure laws relative to those without such laws.

v. State and Year Fixed-Effect Variables

State Fixed-Effect Dummy Variables

The regression includes a total of 50 "state fixed-effect" dummy variables to indicate the state where the respondent lived at the time of interview. Because this regression uses pooled-cross section data, this variable is expected to control for otherwise uncontrolled static differences across states that might have affected an individual's decision to be tested for HIV.¹⁸⁰

176 Phillips, *supra* note 151, at 405-06; Shira Maguen, et al., *Predictors of HIV Antibody Testing Among Gay, Lesbian, and Bisexual Youth*, 26 J. ADOLESCENT HEALTH 252 (2000); Samet, *supra* note 157; Inungu, *supra* note 138.

177 Every year, the survey questionnaire asked whether any of the following applied: "You have used intravenous drugs in the past year;" "You have been treated for a sexually transmitted or venereal disease in the past year;" "You have given or received money or drugs in exchange for sex in the past year;" "You had anal sex without a condom in the past year." *E.g.*, 2009 Behavioral Risk Factor Surveillance System

Questionnaire, CDC (Dec. 30, 2008), http://www.cdc.gov/brfss/annual_data/pdfs/ques/2009brfss.pdf.

178 Maguen, *supra* note 176; Myers et al., *supra* note 163; Samet, *supra* note 157.

179 Inungu, *supra* note 138; Kellerman, *supra* note 138; Samet, *supra* note 157; Spielberg et al., *Overcoming Barriers*, *supra* note 138; Spielberg et al., *Testing Preferences*, *supra* note 138.

180 See JEFFREY M. WOOLDRIDGE, *INTRODUCTORY ECONOMETRICS: A MODERN APPROACH* 456 (2d ed. 2003); Stock, *supra* note 126, at 356-61.

Year Fixed-Effect Dummy Variables

The regression model includes a set of year dummy variables to indicate the year the participant completed the BRFSS interview. Because this study uses survey data conducted over a seven-year time span—from 2002 to 2009 (except 2007)—there could be factors that changed over time but were constant across states, and that other control variables do not control for.¹⁸¹ These dummy variables are expected to capture such changes that might have occurred over time and affected how at-risk individuals reported their recent HIV testing experiences.

vi. Additional Details About Regression Analysis

Clustered-Robust Analysis

In DID analysis, when the treatment variable changes very little within a cluster over time, within-cluster correlation may harm the reliability of the estimates.¹⁸² In this study, the criminalizing policies of the majority of states did not change much during the period of study. Over nine out of 10 samples (91.8%) lived in states that did not introduce an HIV-specific statute between 2002 and 2009; only 8.2% of samples were collected in states where the state adopted an HIV-specific statute during the period of the study.¹⁸³ In addition, once a state adopted an HIV-specific law, it did not abolish it. As a result, there was not much variation in the status of the law within states. In order to account for within-cluster correlation that might occur in DID, this study clusters errors by states.¹⁸⁴

This study also standardizes all continuous variables to achieve convergence. This study standardizes the variables of age, CDC funding, ADAPs' spending, AIDS rate, and the interaction term created between HIV-specific law and the frequency of media reportage. For the analysis, this study uses R as the programming language.

Robustness Check: Lagged Treatment Variables

This study conducts a robustness check to determine whether the results of

¹⁸¹ Stock, *supra* note 126, at 361-62.

¹⁸² Jeffrey W. Wooldridge, Cluster Sample Methods in Applied Econometrics: An Extended Analysis (June 2006) (unpublished manuscript), available at <http://econ.ucsb.edu/~doug/245a/Papers/Cluster%20Sample%20Methods%20in%20Applied%20Econometrics.pdf>; Marianne Bertrand, Esther Duflo & Sendhil Mullainathan, *How Much Should We Trust Differences-in-Differences Estimates?* (Nat'l Bureau of Econ. Research, Working Paper No. 8841, 2002).

¹⁸³ Out of 11,078 observations, 10,172 were collected in states that did not adopt an HIV-specific statute between 2002 and 2009, and the rest (906) were collected in states that introduced a statute between 2002 and 2009.

¹⁸⁴ In R, the lme4 package is used and the "glmer" function is used.

the original regression are reliable. It is difficult to draw a clear line at when exactly after its adoption the criminal law started to affect people's HIV testing decisions. This study, therefore, conducts a separate regression with one-year lagged treatment variables: (1) the lagged HIV-specific law variable, and (2) the interaction term between this variable and the frequency of media reportage variable. As discussed in Section II.B, the results of this regression are nearly identical to those of the original regression.

Robustness Check: Missing-Value Variables

This study uses binary missing-value variables to prevent the bias that dropping observations with missing values from the dataset could cause. In order to check whether observations with missing values have certain traits that are associated with HIV testing, this study first conducts regressions with observations that have missing values. For each control variable, a binary missing-value variable is created and included in the regression. For example, for income level, the income-level missing-value variable is created to indicate whether income level information is missing for that observation. For observations that have a missing value for income level, this missing-value variable is coded 1. Model 1 and Model 3 of the regression in Table 8 present the results of regression with all observations that have missing values.

If the missing-value variable does not have statistical significance in Model 1 and Model 3, observations that have missing values for that variable are dropped from the samples. Missing-value variables for health condition, health insurance coverage, marital status, employment status, race, and pregnancy status do not have statistical significance in Model 1 and Model 3. Therefore, observations that do not have values for these variables are deleted from the data set, and the second group of regressions—Model 2 and Model 4—is conducted with data that do not have missing values for these variables. As Table 8 presents, the results are similar to those of regression using data that include observations with missing values.

2. Description of Samples

i. About BRFSS

The BRFSS is a state-based system of health surveys that collects information on health risk behaviors, preventive health practices, and health care access.¹⁸⁵ The CDC established the BRFSS in 1984, and currently the CDC

185 For a detailed description of the BRFSS, see *Behavioral Risk Factor Surveillance System: About BRFSS*, CDC, <http://www.cdc.gov/brfss/about/> (last updated Mar. 19, 2013). The BRFSS annual survey data from 1984 to 2010 and questionnaires can be downloaded from the CDC website: *Behavioral Risk Factor Surveillance System: Annual Survey Data*, CDC, http://www.cdc.gov/brfss/annual_data/annual_data.htm (last updated July 11, 2013).

collects data monthly in 50 states and the District of Columbia. Data is collected from a random sample of adults (one per household) through a telephone survey. Over 350,000 adults are interviewed each year, but the interview is not repeated to the same individuals.¹⁸⁶

The BRFSS started to cover HIV-related health behaviors in 1993. Although each year's questionnaire has had different forms of questions, the BRFSS usually contains questions regarding recent HIV testing experiences, with an exception in 2007. Since 1998, the questionnaires have asked whether the participant had been tested for HIV within the last year or their last HIV testing had taken place. The BRFSS also collected information about how survey participants perceive their risks of contracting HIV by asking whether any high-risk situations apply. The BRFSS has also consistently contained a wide range of useful information about the demographic characteristics of survey participants.¹⁸⁷

Because the BRFSS is survey data, it has some inherent limitations. The BRFSS relies on the answers of survey participants, so there could be a reporting bias. Given the strong stigma associated with HIV, there is a possibility of underreporting of HIV testing. There could also be a selection bias, since only those who had telephone numbers participated in the BRFSS. In addition, this study selected samples that have answers to HIV testing experience and the high-risk situation.¹⁸⁸ Given the stigma attached to these high-risk groups, it is possible that a substantial portion of individuals at high risk for HIV infection refused to answer the question or did not faithfully answer the question, and accordingly were not included in the regression analysis.

ii. Samples Used in Regression Analysis

The sample contains a total of 11,078 observations selected from the original BRFSS data sets collected from 2002 to 2009 (with the exception of 2007).¹⁸⁹

186 Because of this collection method, the BRFSS is pooled cross-section data, not panel data. Pooled cross-section uses cross-section data collected at two or more different times. Pooled cross-section is analyzed like a standard cross-section, except that differences occurring across time are considered in the analysis. WOOLDRIDGE, *supra* note 180, at 10.

187 This information includes each participant's state of residence, sex, age, race, marital status, education level, employment status, income level, and current pregnancy status. The survey data also has information about each participant's perception of his or her general health condition, and whether the participant had any kind of health insurance coverage at the time of interview.

188 See *supra* note 177.

189 This study excludes the 2007 survey from the analysis because the survey did not have a question about the recent HIV testing experience of the respondent. This study originally attempted to include observations collected from 1998 to 2001 as well. However, during that period, most states that did not adopt an HIV-specific criminal law did not participate in the BRFSS survey. See *infra*, Appendix 1. Insufficient data during the specific survey period raises the question of how the inclusion of data from this period would affect the outcome. Therefore, this study limited its scope of analysis to from 2002 to 2009, excluding 2007.

This data excluded four observations collected in Hawaii in 2004 because the number of

The selection was made based on whether the observation had answers to the recent HIV testing experience question and the high-risk situation question described above.¹⁹⁰ Model 2 and Model 4 regression use data that do not have missing values for all control variables (except income level and education level). Hence, the sample size reduced to 9,705. For the distribution of samples in terms of state and year, see Appendix 1.

Table 7. Sample Demographic and Socio-economic Characteristics

Variable	Individual characteristics	Frequency	Percent
Sex	Male	5242	47.3
	Female	5836	52.7
Ethnicity	White	7295	65.9
	Black	1425	12.9
	Hispanic	1293	11.7
	Other or multiracial	993	9.0
	Missing	72	0.6
Marital status	Married	2612	23.6
	Not Married	8431	76.1
	Missing	35	0.3
Education level	No education or Kindergarten	13	0.1
	Elementary school	236	2.2
	Some high school	1037	9.4
	High school graduate	2959	26.7
	Some college or technical	3129	28.2
	College graduate	3699	33.4
	Missing	5	0.0
Employment status	Employed until within a year	8236	74.3
	Not employed for over a year	2828	25.6
	Missing	14	0.1
Income level	<19,999	2785	25.1
	20,000-34,999	2405	21.7
	35,000-74,999	2937	26.5
	75,000>	2033	18.4
	Missing	918	8.3
Health insurance coverage	Have any health plan	8649	78.1
	No health plan	2410	21.7
	Missing	19	0.2
Self-perceived health condition	Poor health condition	561	5.1
	Fair or better health	10493	94.7

observations in this state and year is not large enough and could thus cause bias. In all four of these observations, all survey participants answered that they had HIV testing in the past 12 months, for a testing rate of 100%. This rate is clearly an outlier compared to the average testing rate of 42.5% in other states.

¹⁹⁰ See *supra* note 177.

	Missing	24	0.2
Current pregnancy status	Currently pregnant	261	2.4
	Not pregnant now	9585	97.6
	Missing	1232	11.1
Reported age in years	Min.	Max.	Mean
	18	64	35.66

Among 11,078 respondents, 4,428 (40%) reported that they had been tested for HIV in the past 12 months for a purpose other than blood donation. 575 respondents (5.2%) lived in the states that had HIV-specific criminal statutes at the time of interview. 906 survey participants (8.2%) lived in states that adopted an HIV-specific statute between 2002 and 2009 (treatment states).

B. Results

The results of the regression in all four specifications show that the HIV-specific law variable does not have a statistically significant impact on the outcome variable. On the other hand, the interaction between the HIV-specific law variable and the media reporting frequency is statistically significant in all models. For every unit of increase in the media reporting of HIV criminalization, a 7% to 9% decrease of the HIV testing rate is expected in states with HIV-specific statutes, all other factors held constant.

Table 8. Regression Results

Dependent variable: Having been tested for HIV in the past 12 months				
Independent variables	Non-lagged treatment variables		Lagged-treatment variables	
	With missing variables (Model 1)	Without missing variables (Model 2)	With missing variables (Model 3)	Without missing variables (Model 4)
HIV-specific law	0.11 (0.17)	0.09 (0.18)	0.16 (0.20)	0.08(0.21)
HIV-specific law*Frequency of media reportage	-0.08(0.03)***	-0.07(0.03)**	-0.09(0.03)***	-0.07(0.03)**
Age	-0.37(0.03)***	-0.35(0.03)***	-0.37(0.03)***	-0.35(0.03)***
Sex	0.31(0.05)***	0.30(0.05)***	0.31(0.05)***	0.30(0.05)***
Poor health condition	0.15(0.10)	0.21(0.11)*	0.15(0.10)	0.21(0.11)*

Health condition missing	0.17(0.43)		0.17(0.43)	
Have health insurance	0.28(0.05)***	0.29(0.06)***	0.28(0.05)***	0.29(0.06)***
Health insurance missing	0.53(0.50)		0.52(0.50)	
Married	-0.71(0.05)***	-0.72(0.06)***	-0.71(0.05)***	-0.73(0.06)***
Married missing	-0.43(0.37)		-0.42(0.37)	
Education level 1 4	-0.09(0.08)	-0.13(0.08)	-0.09(0.08)	-0.14(0.08)
Education level 2 4	0.01(0.06)	0.00(0.06)	0.01(0.06)	0.00(0.06)
Education level 3 4	0.02(0.06)	0.00(0.06)	0.02(0.06)	0.01(0.06)
Education level missing	2.16(1.16)*	2.15(1.16)*	2.17(1.16)*	2.15(1.16)*
Employed	-0.13(0.05)**	-0.12(0.06)**	-0.13(0.05)**	-0.12(0.06)**
Employment missing	-0.43(0.62)		-0.43(0.62)	
Income level 1 4	0.22(0.08)***	0.22(0.08)***	0.22(0.08)***	0.22(0.08)***
Income level 2 4	0.09(0.07)	0.04(0.08)	0.09(0.07)	0.04(0.08)
Income level 3 4	0.02(0.07)	0.01(0.07)	0.02(0.07)	0.01(0.07)
Income level missing	0.25(0.09)***	0.18(0.10)*	0.25(0.09)***	0.18(0.10)*
Race: White	-0.24(0.08)***	-0.21(0.08)***	-0.24(0.08)***	-0.21(0.08)***
Race: Black	0.16(0.10)*	0.16(0.10)	0.16(0.10)	0.16(0.10)
Race: Hispanic	-0.05(0.09)	-0.02(0.10)	-0.05(0.09)	-0.02(0.10)
Race missing	0.35(0.25)		0.34(0.26)	
Pregnancy	1.21(0.15)***	1.22(0.15)***	1.21(0.15)***	1.22(0.15)***
Pregnancy	0.00(0.08)		0.00(0.08)	

missing				
AIDS rate	-0.10(0.30)	-0.07(0.32)	-0.09(0.30)	-0.05(0.32)
CDC funding	0.12(0.13)	0.05(0.14)	0.12(0.13)	0.05(0.14)
ADAPs' spending	-0.03(0.04)	-0.01(0.04)	-0.03(0.04)	-0.02(0.04)

Standard errors in parenthesis

*significance at 10%, **significance at 5%, ***significance at 1%

(Model 1) AIC: 13962, BIC: 14372, logLik: -6925

(Model 2) AIC: 12417, BIC: 12776, logLik: -6158

(Model 3) AIC: 13962, BIC: 14372, logLik: -6925

(Model 4) AIC: 12416, BIC: 12775, logLik: -6158

C. Discussion of Results

1. Regarding HIV-Specific Laws' Influence on HIV Testing

In all specifications, the HIV-specific law variable has no statistically significant impact on HIV testing decisions, holding other conditions constant. State of residence is a broad proxy that reflects HIV-specific statutes' direct and indirect influence on HIV testing decisions concerning at-risk individuals.¹⁹¹ Thus, failures to reject the null hypothesis with this state-of-residence variable do not explain the dynamics of the interaction between criminal law and HIV testing.

The findings of this study are noteworthy, however, because they are in line with many other empirical studies that have found that HIV testing policies do not affect at-risk individuals' HIV testing decisions as much as critics of the policies assume.¹⁹² For example, in one study, researchers examined the number of HIV tests conducted from 1993 to 2000 in one state that adopted mandatory reporting of HIV in 1998. The results showed that HIV testing frequencies did not decrease after introducing the mandatory reporting of HIV infection.¹⁹³ In another empirical study, a group of researchers used HIV Testing Survey (HITS) data collected in California to test at-risk individuals' knowledge of and behavior concerning HIV testing.¹⁹⁴ Contrary to what public health experts thought, this

191 Burris et al., *supra* note 7, at 493.

192 Klitzman et al., *supra* note 119, at 46-47, 49, 50-51.

193 Gayatwi C. Jayaraman et al., *Mandatory Reporting of HIV Infection and Opt-out Prenatal Screening for HIV Infection: Effect on Testing Rates*, 168 CANADIAN MED. ASS'N J. 679 (2003).

194 Sandra Schwarcz et al., *Does Name-Based HIV Reporting Deter High-Risk Persons from HIV Testing? Results from San Francisco*, 35 J. AIDS 93 (2004).

study found that participants who thought California had name-based HIV reporting were more likely to have been tested for HIV infection, a result possibly attributed to participants' poor understanding of reporting regulations; in fact, only 5% of survey participants knew that California had a name-based HIV reporting system, and only 6% were aware of the recent changes in California's HIV reporting policy.¹⁹⁵

An empirical study of HIV-specific laws' influence on sexual behavior of at-risk individuals also mentions that an HIV-specific law is likely to have a marginal influence on HIV testing. Scott Burris and colleagues found that residence in two different states with different HIV criminal law was not associated with reduced reporting of risky sexual activity. The research also found that residence in different states was not correlated with different beliefs about HIV criminal law.¹⁹⁶ The researchers concluded that their results do not "lend support" to the claim that HIV-specific laws influence people at high risk of HIV infection to shun public health services.¹⁹⁷

The empirical study presented in this Note does not alone provide sufficient information about how HIV-specific statutes affect HIV testing decisions, as such statutes could have had a wide range of direct and indirect influences on HIV testing. The study does show, however, that these influences cannot be said to cause at-risk individuals to report different testing behavior.

2. The Media's Intermediary Role between Criminalization and HIV Testing Behavior

The coefficient of the interaction term between the HIV-specific law variable and the media reportage frequency variable is statistically significant in all model specifications; increased media reporting of criminalization in states with HIV-specific statutes is correlated with a fewer number of people who reported having been tested for HIV in the past year.

This study used the number of newspaper reports on criminalization of behavior that exposes others to HIV as a rough proxy for media coverage intensity. The study's frequency count includes newspaper reports on criminal prosecutions brought against HIV-positive individuals for knowingly exposing others to HIV as well as on debates surrounding the adoption of a statute criminalizing HIV exposure.¹⁹⁸ This proxy does not reveal why and how such media coverage affects the law's influence, but it does show that a factor associated with this frequency affects the law's influence on HIV testing.

News reports, together with introduction of an HIV-specific criminal statute, may have heightened people's awareness of the criminal law, which led to less HIV testing. Frequent media coverage of criminalization may have alerted at-risk

¹⁹⁵ *Id.* at 94-95.

¹⁹⁶ Burris et al., *supra* note 7, at 502.

¹⁹⁷ *Id.* at 512.

¹⁹⁸ See *supra* Subsection II.A.3.

individuals that positive HIV test results could be used against them to prove criminal liability for knowingly exposing others to HIV.

Alternatively, the adoption of a criminal statute and the robust media reporting of HIV exposure's criminalization could have deepened social hostility against HIV-positive individuals. HIV-specific statutes could have singled out the HIV-positive population from the rest of society and stigmatize this population as a dangerous group requiring special attention and social regulation.¹⁹⁹ In addition, the media could have sensationalized high-profile cases involving extremely condemnable HIV-positive individuals who intentionally infected others with HIV.²⁰⁰ The prevalence of this negative sentiment could have deterred at-risk individuals from utilizing public health services, including HIV testing.

These possibilities are highly likely, based on what is already known about HIV. Yet, these theories are not evidence-based. In order to provide recommendations for criminal policy that supports public health goals, it is necessary to gather further evidence on the interactions between the media, criminalization, and HIV testing.

III. GUIDANCE FOR FUTURE RESEARCH EFFORTS

This Note finds that criminalization of behavior that exposes others to HIV might not exert as significant an influence on HIV prevention as claimed by public health experts and advocacy groups.²⁰¹ These groups argue that at-risk individuals might abstain from HIV testing based on fear of punitive sanctions for knowingly exposing others to HIV.²⁰² In fact, empirical studies have identified the fear of criminal prosecution as a possible deterrent to HIV testing.²⁰³ However, the results of the empirical analysis in this Note failed to find support for this claim; residence in states with HIV-specific criminal statutes was not associated with a fewer number of at-risk individuals who reported that they had been tested for HIV in the past year.

199 Elliot, *supra* note 95, at 32; *Verdict on a Virus*, *supra* note 111, at 13; *HIVMA Urges Repeal of HIV-Specific Criminal Statutes*, HIV MED. ASS'N 2 (2012), <http://www.hivma.org/uploadedFiles/HIVMA/FINAL%20HIVMA%20Policy%20Statement%20on%20HIV%20Criminalization.pdf>; Understanding State Dep'ts of Health and Corrections Collaboration, *A Summary of Survey Findings — Part II and Strategic Guidance Towards Ending Criminalization-Related Stigma and Discrimination*, NAT'L ALLIANCE STATE & TERRITORIAL AIDS DIRECTORS (Dec. 2011), http://www.nastad.org/HIVC/decriminalization_findings.pdf.

200 Jürgens et al., *supra* note 123, at 166; *Verdict on a Virus*, *supra* note 111, at 24-26; Elliot, *supra* note 95, at 23-24.

201 See *supra* Subsection I.E.2.

202 See, e.g., Ruth Lowbury, *Criminal Prosecution for HIV Transmission*, 333 BRIT. MED. J. 666, 666-67 (2006).

203 For empirical studies, see Dodds et al., *supra* note 7, at 140-41 (noting that some participants in a focus group cited criminalization as a deterrence to HIV testing) and Klitzman et al., *supra* note 119, at 49-50 (noting that some survey participants expressed fear that criminalization of behaviors exposing others to HIV might deter people from being tested).

However, because the regression shows another contradicting result, the conclusion that criminal laws might not exert a significant influence on HIV prevention should be made with caution. Regression analysis also indicates that the adoption of a new HIV-specific criminal law together with intense media coverage of this criminalization is associated with reduced HIV testing in at-risk individuals. These results indicate that HIV testing decisions could be affected if the media actively broadcasts about criminal law. The specific reason why such heavy broadcasting of the law is associated with the decreased number of people being tested for HIV is unknown. It is possible that media reporting of HIV-specific statutes increases stigma against HIV-positive individuals, making at-risk individuals hesitant to come forward for HIV testing. Media reporting could also increase at-risk individuals' awareness of the law, thereby making them fearful that their HIV test results could be used against them in criminal courts. In any event, this result proves that the concerns critics of criminalization of HIV-exposing behavior have raised may have a point. When coupled with certain factors, criminal law can have a substantially negative impact on HIV testing of at-risk individuals.

To provide a detailed recommendation for how criminal law should be changed to prevent such a negative impact on HIV testing, further research exploring the dynamic of HIV criminalization, media coverage, and HIV testing is necessary. To understand the dynamic, future research should investigate individuals' perceptions of the law and the impact of these perceptions on HIV testing decisions. In addition, measurement of the relationship between stigma, HIV criminalization, media coverage, and HIV testing is crucial to improving the criminal law.