

Faculty Publications

---

2019

## What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation

Sharona Hoffman

Case Western Reserve University School of Law, [sharona.hoffman@case.edu](mailto:sharona.hoffman@case.edu)

Follow this and additional works at: [https://scholarlycommons.law.case.edu/faculty\\_publications](https://scholarlycommons.law.case.edu/faculty_publications)

 Part of the [Health Law and Policy Commons](#)

---

### Repository Citation

Hoffman, Sharona, "What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation" (2019). *Faculty Publications*. 2035.

[https://scholarlycommons.law.case.edu/faculty\\_publications/2035](https://scholarlycommons.law.case.edu/faculty_publications/2035)

This Article is brought to you for free and open access by Case Western Reserve University School of Law Scholarly Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Case Western Reserve University School of Law Scholarly Commons.

## **What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation**

**Sharona Hoffman\***

*The ever-growing phenomenon of predictive health analytics is generating significant excitement, hope for improved health outcomes, and potential for new revenues. Researchers are developing algorithms to predict suicide, heart disease, stroke, diabetes, cognitive decline, future opioid abuse, and other ailments. The researchers include not only medical experts, but also commercial enterprises such as Facebook and LexisNexis, who may profit from the work considerably. This Article focuses on long-term disease predictions (predictions regarding future illnesses), which have received surprisingly little attention in the legal and ethical literature. It compares the robust academic and policy debates and legal interventions that followed the emergence of genetic testing to the relatively anemic reaction to predictions produced by artificial intelligence and other predictive methods. The paper argues that like genetic testing, predictive health analytics raise significant concerns about psychological harm, privacy breaches, discrimination, and the meaning and accuracy of predictions. Consequently, as alluring as the new predictive technologies are, they require careful consideration and thoughtful safeguards. These include changes to the HIPAA Privacy and Security Rules and the Americans with Disabilities Act, careful oversight mechanisms, and self-regulation by health care providers. Ignoring the hazards of long-term predictive health analytics and failing to provide data subjects with appropriate rights and protections would be a grave mistake.*

---

\* Edgar A. Hahn Professor of Law and Professor of Bioethics, Co-Director of Law-Medicine Center, Case Western Reserve University School of Law; B.A., Wellesley College; J.D., Harvard Law School; LL.M. in Health Law, University of Houston; S.J.D. in Health Law, Case Western Reserve University. Author of *Electronic Health Records and Medical Big Data: Law and Policy* (Cambridge University Press 2016). A huge thank you to Computer Science Professor Andy Podgurski for everything he taught me about predictive health analytics and for his astute comments on prior drafts. I also thank Melissa Vogley for her skillful research assistance.

## I. INTRODUCTION

The ever-growing phenomenon of predictive health analytics is generating significant excitement, hope for improved health outcomes, and potential for new revenues.<sup>1</sup> Researchers are developing algorithms to predict suicide, heart disease, stroke, diabetes, cognitive decline, future opioid abuse, and other ailments.<sup>2</sup> In 2017 the Society of Actuaries found that ninety-three percent of health care and health insurance executives that it surveyed believed that predictive analytics is important to their future success.<sup>3</sup> Indeed, experts forecast that predictive health analytics will be a commonplace medical tool in the near future.<sup>4</sup>

Health care providers can use predictive health analytics for treatment purposes in the short term.<sup>5</sup> For example, predictive health analytics can help physicians identify patients who are at risk for hospital re-admission because of complications.<sup>6</sup> This Article focuses, however, on health analytics that predict health problems in the more distant future, which I call “long-term predictive health analytics.” For instance, scientists are developing techniques to forecast conditions such as heart disease or cognitive decline that are years or decades away.<sup>7</sup>

In some instances, such forecasts can be medically beneficial because clinicians can commence early screening of affected individuals and implement preventive interventions.<sup>8</sup> In the case of

---

<sup>1</sup> Jennifer Bresnick, *10 High-Value Use Cases for Predictive Analytics in Healthcare*, HEALTH IT ANALYTICS (Sept. 4, 2018), <https://healthitanalytics.com/news/10-high-value-use-cases-for-predictive-analytics-in-healthcare>.

<sup>2</sup> See *infra* Parts II.B & C.

<sup>3</sup> Society of Actuaries, *Predictive Analytics in Healthcare Trend Forecast 2*, 4 (2017), <https://www.soa.org/Files/programs/predictive-analytics/2017-health-care-trend.pdf>. The survey included 223 participants. *Id.* at 2.

<sup>4</sup> Eric J. Topol, *High-Performance Medicine: the Convergence of Human and Artificial Intelligence*, 25 NATURE MED. 44, (2019).

<sup>5</sup> I. Glenn Cohen et al., *The Legal and Ethical Concerns that Arise from Using Complex Predictive Analytics in Healthcare*, 33 HEALTH AFF. 1139, 1140 (2014) (explaining that “it has become possible to apply predictive analytics to health care”).

<sup>6</sup> *Id.*

<sup>7</sup> See *infra* Parts II.B. & C.

<sup>8</sup> Bresnick, *supra* note 1.

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

heart disease, these might include drugs, exercise, and improved diet.<sup>9</sup>

At the same time, predictive health analytics can also be potentially harmful.<sup>10</sup> Individuals who are identified as being at high risk of developing future health problems such as cognitive decline or opioid addiction may suffer psychological distress, privacy violations (if the information is circulated to unauthorized third parties), discrimination, and other harms.<sup>11</sup> One scholar worries that people labelled as being at high risk for suicide will be treated differently by their physicians.<sup>12</sup> Physicians might discontinue beneficial medications for fear that they will exacerbate the suicide risk, unnecessarily send police to patients' homes, forcibly hospitalize individuals, or relate to them in a demeaning, dehumanizing way.<sup>13</sup>

Moreover, predictive health analytics outcomes can be erroneous for a variety of reasons.<sup>14</sup> Thus, individuals may endure serious adverse consequences based on mistaken predictions, when in truth there is no evidence that they are at risk of developing the health problems at issue.

This Article argues that we are doing alarmingly little to identify and address the ethical and legal implications of long-term predictive health analytics. This is in stark contrast to policy-makers' thoughtful approach to the emergence of genetic testing several decades ago.<sup>15</sup>

The Article highlights the discrepancy between society's relatively cautious approach to genetic testing and its more cavalier approach to predictive analytics. It argues that scientists must carefully consider the benefits and risks of predictive health analytics and implement safeguards to address their hazards.

Data subjects should enjoy specified rights that give them a degree of control over their data, including predicted health outcomes. They should have an expanded right to consent to

---

<sup>9</sup> *Strategies to Prevent Heart Disease*, MAYO CLINIC (Jan. 9, 2019), <https://www.mayoclinic.org/diseases-conditions/heart-disease/in-depth/heart-disease-prevention/art-20046502>.

<sup>10</sup> *See infra* Part IV.

<sup>11</sup> *Id.*

<sup>12</sup> Mason Marks, *Artificial Intelligence Based Suicide Prediction*, YALE J. HEALTH POL'Y L. & ETHICS \_\_ (forthcoming)

<sup>13</sup> *Id.* at \_\_.

<sup>14</sup> *See infra* Part IV.

<sup>15</sup> *See infra* Part III.

disclosure of their health information, a right to discover who has seen their health data, and a right to sue for privacy breaches that harm them and for discrimination based on disease predictions.<sup>16</sup> The scientific community should also develop an oversight mechanism to safeguard the quality of predictive models.<sup>17</sup>

The remainder of the Article proceeds as follows. Part II describes long-term predictive health analytics and illustrates the work that scientists are conducting in this area. Part III analyzes the precedent of genetic testing, focusing on the concerns that it raised and the measures that policy-makers implemented to address them. Part IV examines the risks of long-term predictive health analytics. Part V develops preliminary recommendations for responsive legal and policy changes. Part VI concludes.

## II. LONG-TERM PREDICTIVE HEALTH ANALYTICS

### A. Predictive Health Analytics Defined

Predictive analytics can be defined as “the analysis of large data sets to discover patterns and use those patterns to forecast or predict the likelihood of future events.”<sup>18</sup> Experts conduct this analysis using computer algorithms.<sup>19</sup> An algorithm is a precise step-by-step process that leaves nothing to guesswork or intuition.<sup>20</sup> Learning algorithms train predictive models using training sets comprised of sample input and output values.<sup>21</sup> Some analysts use the term “predictive modeling” which can be defined as “the process of developing a mathematical tool or model that generates an accurate prediction.”<sup>22</sup> Researchers often use the terms “learning algorithm” and “predictive model” interchangeably, although the term “predictive model” suggests a representation of knowledge that is

---

<sup>16</sup> See *infra* Part V.A.1 & 2.

<sup>17</sup> See *infra* Part V.B.1.

<sup>18</sup> David Crockett et al., *What is Data Mining in Healthcare?* HEALTHCATALYST (2017) <https://www.healthcatalyst.com/wp-content/uploads/2014/06/What-is-data-mining-in-healthcare.pdf>.

<sup>19</sup> I. Glenn Cohen et al., *supra* note 5, at 1139; Nicholson Price, *Regulating Black-Box Medicine*, 116 MICH. L. REV. 421, 425-26 (2017) (discussing the nature of medical algorithms).

<sup>20</sup> Deven R. Desai & Joshua A. Kroll, *Trust but Verify: A Guide to Algorithms and the Law*, 31 HARV. J.L. & TECH. 1, 23 (2017).

<sup>21</sup> SHAI SHALEV-SHWARTZ & SHAI BEN DAVID, UNDERSTANDING MACHINE LEARNING 13-14 (2014) (discussing “the statistical learning framework”).

<sup>22</sup> MAX KUHN & KJELL JOHNSON, APPLIED PREDICTIVE MODELING 2 (2013).

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

created by an algorithm.<sup>23</sup> Predictive analysis is based on techniques from three closely related areas of research: statistical inference, data mining, and machine learning.

Statistical inference involves analyzing a dataset and, based on this sample, inferring properties of a larger population and characterizing uncertainties about them.<sup>24</sup> Data mining is the process of using algorithms to examine “big data” from sources such as databases or the Internet in order to unearth hidden knowledge or patterns.<sup>25</sup> “Big data” can be defined as data that is of high volume, variety, and velocity, the latter of which refers to the speed with which it is generated.<sup>26</sup> In medicine, big data can come from a myriad of sources, including patients, health care providers, insurers, manufacturers, the government, and even mobile devices such as smartphones and wearables.<sup>27</sup>

Machine learning refers to methods that enable computers to “automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision-making under uncertainty.”<sup>28</sup> Scientists train computers to do analytical work by feeding them information, such as patients’ medical records.<sup>29</sup> For example, scientists might show computers a large number of tumor images with indications as to which ones are

---

<sup>23</sup> See *supra* notes 21-22 and accompanying text.

<sup>24</sup> WILLIAM L. HAYS, STATISTICS 1 (4<sup>TH</sup> ed. 1988) (describing statistical inference as a process of analysis that enables one to “make general statements about the large body of potential observations, of which the data collected represents but a sample”); *Statistical Inference*, OXFORD LIVING DICTIONARIES, [https://en.oxforddictionaries.com/definition/us/statistical\\_inference](https://en.oxforddictionaries.com/definition/us/statistical_inference) (last visited Feb. 25, 2019) (defining statistical inference as “[t]he theory, methods, and practice of forming judgments about the parameters of a population and the reliability of statistical relationships, typically on the basis of random sampling”).

<sup>25</sup> JIAWEI HAN ET AL., DATA MINING: CONCEPTS AND TECHNIQUES 8 (3<sup>RD</sup> ED. 2012).

<sup>26</sup> SHARONA HOFFMAN, ELECTRONIC HEALTH RECORDS AND MEDICAL BIG DATA: LAW AND POLICY 111 (2016).

<sup>27</sup> Nathan Cortez, *Substantiating Big Data in Health Care*, 14 I/S: J.L. & POL’Y FOR INFO. SOC’Y 61, 63-65 (2017) (discussing the breadth of big data sources).

<sup>28</sup> KEVIN P. MURPHY, MACHINE LEARNING: A PROBABILISTIC PERSPECTIVE 1 (2012). See also, David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn about Machine Learning*, 51 U.C. DAVIS L. REV. 653, 671 (2017) (“Fundamentally, machine learning refers to an automated process of discovering correlations (sometimes alternatively referred to as relationships or patterns) between variables in a dataset, often to make predictions or estimates of some outcome.”).

<sup>29</sup> See Niha Beig et al., *Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas*, 290 RADIOLOGY 783, 784 (2019) (relating that a “machine classifier was trained on a cohort of 145 patients”).

*What Genetic Testing Teaches About Long-Term Predictive Health  
Analytics Regulation*

cancerous and which ones are not.<sup>30</sup> The computers then learn to differentiate between benign and malignant tumors based on patterns in the tumor x-rays or scans, so that they can identify cancerous tumors when shown new images.<sup>31</sup>

A well-known type of machine learning is deep learning, which allows computers “to learn from experience and understand the world in terms of a hierarchy of concepts, with each concept defined through its relation to simpler concepts.”<sup>32</sup> Thus, computers gather knowledge from experience and learn more complex concepts by building on simpler concepts.<sup>33</sup> Many readers will be familiar with the general term “artificial intelligence,” which refers to computers’ ability to mimic human behavior and learn.<sup>34</sup>

Predictive models are valuable for physicians, researchers, and policy makers.<sup>35</sup> They can help public health officials identify those who are at highest risk of developing a disease so they can implement preventive interventions for them.<sup>36</sup> In the clinical setting, predictive models may discern which patients are likely to have poor and successful treatment outcomes so physicians can tailor their medical decisions accordingly. Predictive analytics may also help identify high-risk individuals whom doctors should aggressively screen for particular diseases.<sup>37</sup>

In addition, predictive health analytics can generate projections regarding the health problems that will plague individuals in the future.<sup>38</sup> These long-term forecasts are the subject of this Article. Such predictions can be beneficial to patients if physicians can offer medical interventions to prevent or detect the condition at issue at a very early stage. However, such predictions can also render the data subject vulnerable to adverse psychological consequences, discrimination, and other harms.<sup>39</sup>

---

<sup>30</sup> *Id.*

<sup>31</sup> *Id.* at 792.

<sup>32</sup> IAN GOODFELLOW ET AL., DEEP LEARNING 2 (2016).

<sup>33</sup> *Id.*

<sup>34</sup> *Id.*

<sup>35</sup> EWOUT W. STEYERBERG, CLINICAL PREDICTION MODELS 1-3, 11 (2009).

<sup>36</sup> *Id.* at 11-12.

<sup>37</sup> *Id.*

<sup>38</sup> See *infra* Parts II.B. and C.

<sup>39</sup> See *infra* Part IV.

## **B. Long-Term Predictive Health Analytics Examples**

Scientists are working hard to identify physical and behavioral clues to individuals' future health status. Many studies focus on the question of whether there are traits, habits, or other indicators that signal that a person is vulnerable to particular diseases in the future.

Medical researchers are investigating biomarkers that can help them discern disease risks. A “biomarker” is a “biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease.”<sup>40</sup> For example, researchers reported in 2014 that in their study, people with lower levels of ten phospholipids in their blood were at higher risk of suffering cognitive impairments at the time of the blood draw or within a few years.<sup>41</sup>

A 2018 study found that retinopathy was associated with higher rates of cognitive decline over the next twenty years.<sup>42</sup> Retinopathy is a disease that involves the small retinal blood vessels in the eye.<sup>43</sup>

Human eyes can also reveal information about cardiovascular risks. As reported in 2018, researchers from Google and its health-tech subsidiary, Verily, used machine learning to analyze eye scans and medical data from nearly 300,000 patients in order to develop an algorithm that can predict individuals' risk of heart disease.<sup>44</sup>

---

<sup>40</sup> *NCI Dictionary of Cancer Terms*, NAT'L CANCER INST., [www.cancer.gov/publications/dictionaries/cancer-terms?cdrid=45618](http://www.cancer.gov/publications/dictionaries/cancer-terms?cdrid=45618) (last visited Mar. 7, 2019).

<sup>41</sup> Alison Abbott, *Biomarkers Could Predict Alzheimer's Before It Starts*, NATURE, Mar. 9, 2014, available at [www.nature.com/news/biomarkers-could-predict-alzheimer-s-before-it-starts-1.14834](http://www.nature.com/news/biomarkers-could-predict-alzheimer-s-before-it-starts-1.14834).

<sup>42</sup> Jennifer A. Deal et al., *Retinal Signs and 20-year Cognitive Decline in the Atherosclerosis Risk in Communities Study*, 90 NEUROLOGY e1158, e1158 (2018). The study involved 12,317 men and women who were 50 to 73 years of age when they were first examined.

<sup>43</sup> Janet M. Torpy et al., *Retinopathy*, 298 JAMA. 944, 944 (2007).

<sup>44</sup> Ryan Poplin et al., *Prediction of Cardiovascular Risk Factors from Retinal Fundus Photographs via Deep Learning*, 2 NATURE BIOMEDICAL ENGINEERING 158, 158 (2018); James Vincent, *Google's New AI Algorithm Predicts Heart Disease by Looking at Your Eyes*, THE VERGE (Feb. 19, 2018) (“As with all deep learning analysis, neural networks were then used to mine this information for patterns, learning to associate telltale signs in the eye scans with the metrics needed to predict cardiovascular risk (e.g., age and blood pressure).”).



*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

IBM researchers “identified an automated machine-learning speech classifier” that could predict psychosis based on the speech patterns of high-risk patients.<sup>45</sup> The technique relied on indicators such as less semantic coherence and diminished use of possessive pronouns, and it reportedly achieved an eighty-three percent accuracy rate.<sup>46</sup>

Electronic documentation has been particularly helpful for purposes of health predictions. In 2014, IBM announced that it had analyzed electronic health records from Virginia’s Carilion Clinic and been able to identify 8,500 patients who were at risk of heart failure.<sup>47</sup> Scientists have also been able to use analysis of electronic health records and medical claims data to predict which individuals will develop depression or diabetes-related problems up to a year in advance.<sup>48</sup> The VA has launched a program called “VA Reach Vet” by which it uses a predictive model to analyze veterans’ electronic health records to identify individuals at high risk of suicide.<sup>49</sup> The Society of Actuaries used the Health Care Cost Institute database, containing a seven-year record of insurance claims from forty-seven million individuals, to predict which patients would have the highest costs.<sup>50</sup> It found that the most telling factor is prior cost history, followed by age, gender, and prescription drug coverage.<sup>51</sup>

---

<sup>45</sup> Cheryl M. Corcoran et al., *Prediction of Psychosis Across Protocols and Risk Cohorts Using Automated Language Analysis*, 17 WORLD PSYCHIATRY 67, 67 (2018).

<sup>46</sup> *Id.*

<sup>47</sup> *IBM Predictive Analytics to Detect Patients at Risk for Heart Failure*, IBM NEWS RELEASE, Feb. 19, 2014, available at <http://www-03.ibm.com/press/us/en/pressrelease/43231.wss>; Mohana Ravindranath, *IBM Used Predictive Analytics to Find Patients at Risk of Heart Failure*, WASH. POST, Feb. 20, 2014, [www.washingtonpost.com/business/on-it/ibm-used-predictive-analytics-to-find-patients-at-risk-of-heart-failure/2014/02/20/9b0ddb3c-9a47-11e3-b88d-f36c07223d88\\_story.html](http://www.washingtonpost.com/business/on-it/ibm-used-predictive-analytics-to-find-patients-at-risk-of-heart-failure/2014/02/20/9b0ddb3c-9a47-11e3-b88d-f36c07223d88_story.html)

<sup>48</sup> Arthur Allen, *Big Brother Is Watching Your Waist*, POLITICO (July 21, 2014), [www.politico.com/story/2014/07/data-mining-health-care-109153](http://www.politico.com/story/2014/07/data-mining-health-care-109153).

<sup>49</sup> *VA REACH VET Initiative Helps Save Veterans Lives: Program Signals When More Help Is Needed for At-Risk Veterans*, U.S. DEP’T VETERANS AFFAIRS (Apr. 3, 2017), <https://www.va.gov/opa/pressrel/includes/viewpdf.cfm?id=2878>.

<sup>50</sup> Brian Hartman et al., *Predicting High-Cost Members in the HCCI Database* (July 20, 2018) (unpublished manuscript), available at [https://hartman.byu.edu/files/HartmanOwenGibbs\\_HighCostClaims.pdf](https://hartman.byu.edu/files/HartmanOwenGibbs_HighCostClaims.pdf).

<sup>51</sup> *Id.* at 4.

### **C. Non-Traditional Data Sources for Predictive Health Analytics**

Social media has become an increasingly common source of data used for predictive health analytics. Researchers recently reported that they used an algorithm to analyze Facebook data from close to 1200 consenting users and identified linguistic signals that could predict depression.<sup>52</sup>

Facebook itself has joined the fray of predictive health analytics. Its software now monitors users' posts to identify those with suicidal intent, and an algorithm assigns a risk score ranging from zero to one.<sup>53</sup> The algorithm interprets phrases such as "Are you okay?" paired with "Goodbye" and "Please don't do this" as clues that someone is in distress.<sup>54</sup> In cases it assesses as severe, Facebook contacts the police, as it did at least 3,500 times in 2018.<sup>55</sup> Unfortunately, police officers who are poorly trained or inexperienced may mishandle such "wellness checks," exacerbating the situation and in extreme cases, using deadly force against individuals with mental illness.<sup>56</sup>

Analysts are turning to other nontraditional data sources as well. For example, several years ago, Carolinas Healthcare (now Atrium Health) purchased consumer information from data brokers<sup>57</sup> in an

---

<sup>52</sup> Johannes C. Eichstaedt et al., *Facebook Language Predicts Depression in Medical Records*, 115 PNAS 11203, 11203, 11207 (2018).

<sup>53</sup> Martin Kaste, *Facebook Increasingly Reliant on A.I. To Predict Suicide Risk*, NPR (Nov. 17, 2018), <https://www.npr.org/2018/11/17/668408122/facebook-increasingly-reliant-on-a-i-to-predict-suicide-risk>; Benjamin Goggin, *Inside Facebook's Suicide Algorithm: Here's How the Company Uses Artificial Intelligence to Predict Your Mental State from Your Posts*, BUS. INSIDER (Jan. 6, 2019), <https://www.businessinsider.com/facebook-is-using-ai-to-try-to-predict-if-youre-suicidal-2018-12>.

<sup>54</sup> Mason Marks, *Suicide Prediction Technology Is Revolutionary. It Badly Needs Oversight*, WASH. POST, Dec. 20, 2018, [https://www.washingtonpost.com/outlook/suicide-prediction-technology-is-revolutionary-it-badly-needs-oversight/2018/12/20/214d2532-fd6b-11e8-ad40-cdfd0e0dd65a\\_story.html?utm\\_term=.951089ae3f76](https://www.washingtonpost.com/outlook/suicide-prediction-technology-is-revolutionary-it-badly-needs-oversight/2018/12/20/214d2532-fd6b-11e8-ad40-cdfd0e0dd65a_story.html?utm_term=.951089ae3f76).

<sup>55</sup> Kaste, *supra* note 53.

<sup>56</sup> Marks, *supra* note 12, at \_\_\_\_.

<sup>57</sup> Data brokers are "companies that collect information, including personal information about consumers, from a wide variety of sources for the purpose of reselling such information to their customers for various purposes . . ." FEDERAL TRADE COMMISSION, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE 68 (2012), available at <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade->

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

effort to use algorithms that identify high-risk patients.<sup>58</sup> Information garnered from credit card purchasing records or grocery loyalty cards can indicate whether individuals are buying healthy food, smoking, refilling their prescriptions, and have gym memberships.<sup>59</sup> These data in turn can predict the likelihood that someone will have a severe asthma attack or a heart attack.<sup>60</sup>

The University of Pittsburgh Medical Center has used patient demographic and household information to predict health risks.<sup>61</sup> It concluded that people who do not reside with children in the home and earn less than \$50,000 annually are more likely to visit emergency rooms rather than to make an appointment with a private doctor, which is a much cheaper way to obtain appropriate treatments for many conditions.<sup>62</sup> Likewise, individuals without a car may not be receiving adequate medical care.<sup>63</sup> Healthcare systems assert that they use such information in order to implement preventive and corrective interventions for patients.<sup>64</sup> However, skeptics have questioned their true motivations, suspecting that cost savings are at the heart of the matter and worrying that data mining practices compromise patient privacy and damage the physician-patient relationship.<sup>65</sup>

#### **D. Predictive Health Analytics as Big Business**

Predictive health analytics have generated business opportunities for enterprising organizations. Companies are reportedly selling “risk scores” to health care providers and insurers to identify patients who are at risk of becoming addicted to or

---

[commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf](#). See also Janine S. Hiller, *Healthy Predictions? Questions for Data Analytics in Health Care*, 53 AM. BUS. L.J. 251, 271-72 (2016) (discussing data brokers and health data streams).

<sup>58</sup> Melanie Hicken, *Big Data: Look Who's Buying Your Personal Information*, CNN MONEY, Sept. 10, 2014, <https://money.cnn.com/gallery/pf/2014/09/07/big-data-personal-information/3.html>; Shannon Pettypiece & Jordan Robertson, *Hospitals, Including Carolinas HealthCare, Using Consumer Purchase Data for Information on Patient Health*, THE CHARLOTTE OBSERVER, June 27, 2014, <https://www.charlotteobserver.com/living/health-family/article9135980.html>.

<sup>59</sup> Pettypiece & Robertson, *supra* note 58.

<sup>60</sup> *Id.*

<sup>61</sup> *Id.*

<sup>62</sup> *Id.*

<sup>63</sup> *Id.*

<sup>64</sup> *Id.*

<sup>65</sup> *Id.* See *supra* Part IV for discussion of concerns raised by predictive analytics.

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

overdosing on opioids.<sup>66</sup> Business giants such as LexisNexis collect data from insurance claims, electronic health records, housing information, and records relating to patients' social and family connections in order to produce risk scores.<sup>67</sup> They do all this without asking patients for permission and are not required to seek consent by law.<sup>68</sup>

Data brokers sell other types of information to health care providers as well. LexisNexis and Acxiom sell assessments of patients based on "criminal records, online purchasing histories, retail loyalty programs and voter registration data."<sup>69</sup> This information is used to identify individuals who are at risk of requiring costly care or readmission to a hospital.<sup>70</sup>

Moreover, data brokers routinely supply predictive health information to parties outside the health care industry as well. They garner data from a myriad of sources such as publicly available records, surveys, shopper loyalty programs, social media, magazine subscription lists, fitness devices, people's Internet searches, and more.<sup>71</sup> They then organize and sell the data, often with personally identifying information, to interested third parties, including marketers.<sup>72</sup> These buyers can use the medical information for

---

<sup>66</sup> Mohana Ravindranath, *How Your Health Information Is Sold and Turned into 'Risk Scores'*, POLITICO (Feb. 3, 2019), <https://www.politico.com/story/2019/02/03/health-risk-scores-opioid-abuse-1139978>. See also, Timothy R. Hylan et al., *Automated Prediction of Risk for Problem Opioid Use in a Primary Care Setting*, 16 J. PAIN 380, 385 (2015) (discussing use of electronic health records to develop "simple risk stratification algorithms to initially alert clinicians to ... patients at higher risk for problem opioid use").

<sup>67</sup> Ravindranath, *supra* note 66.

<sup>68</sup> *Id.* See *supra* notes 142-143 and accompanying text for a discussion of the limitations of the HIPAA Privacy Rule.

<sup>69</sup> Mohana Ravindranath, *Does Your Doctor Need to Know What You Buy on Amazon?*, POLITICO (Oct. 30, 2018), <https://www.politico.com/story/2018/10/30/the-doctor-will-see-through-you-now-893437>.

<sup>70</sup> *Id.*

<sup>71</sup> Adam Tanner, *Strengthening the Protection of Patient Medical Data*, THE CENTURY FOUNDATION (Jan. 10, 2017), <https://tcf.org/content/report/strengthening-protection-patient-medical-data/?agreed=1>; Sam Thielman, *Your Private Medical Data Is for Sale – and It's Driving a Business Worth Billions*, THE GUARDIAN (Jan. 17, 2017), <https://www.theguardian.com/technology/2017/jan/10/medical-data-multibillion-dollar-business-report-warns>.

<sup>72</sup> Tanner, *supra* note 71.

purposes of predictive analytics, to predict individuals' future behaviors and health needs.<sup>73</sup>

### III. THE PRECEDENT OF GENETIC TESTING

Predictive health analytics are novel and exciting, but they are not the first mechanism used to predict future health problems. A much more familiar and well-established technique is genetic testing, also known as DNA testing.<sup>74</sup> When genetic testing emerged as a prevalent diagnostic and predictive tool, it raised significant ethical, legal, and policy concerns. There is much to be learned from the conversations and interventions that followed.

#### A. Predictive Genetic Testing

In the late 1960s scientists developed the ability to test fetuses for Down's syndrome through a sample of amniotic fluid.<sup>75</sup> Fetal genetic testing became common beginning in the 1970s, and today it is used to screen for Tay-Sachs disease, sickle cell disease, cystic fibrosis, and many other illnesses.<sup>76</sup>

Genetic testing can also analyze disease risks after birth and provide information regarding the likelihood that individuals will develop specific maladies in the future.<sup>77</sup> In 1990, Mary King-

---

<sup>73</sup> Russ Cobb, 2018: *Taking Your Healthcare System Marketing Strategies into the Consumer Age*, BECKER'S HOSPITAL REV. (Dec. 19, 2017), <https://www.beckershospitalreview.com/patient-engagement/2018-taking-your-healthcare-system-marketing-strategies-into-the-consumer-age.html>. Thielman, *supra* note 71.

<sup>74</sup> Wylie Burke, *Genetic Testing*, 347 NEW ENG. J. MED. 1867, 1867 (2002); *Frequently Asked Questions about Genetic Testing*, NAT'L HUMAN GENOME RESEARCH INST., <https://www.genome.gov/19516567/faq-about-genetic-testing/> (last updated Feb. 13, 2019) ("Genetic testing uses laboratory methods to look at your genes, which are the DNA instructions you inherit from your mother and your father. Genetic tests may be used to identify increased risks of health problems, to choose treatments, or to assess responses to treatments."); *Genetic Testing: How It Is Used for Healthcare*, DEP'T OF HEALTH & HUMAN SERVS., <https://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=43> (last updated June 30, 2018).

<sup>75</sup> Glenn E. Palomaki, *Screening for Down's Syndrome*, 333 NEW ENGL. J. MED. 532 (1995).

<sup>76</sup> Nancy Press, *Genetic Testing and Screening*, in FROM BIRTH TO DEATH AND BENCH TO CLINIC: THE HASTINGS CENTER BIOETHICS BRIEFING BOOK FOR JOURNALISTS, POLICYMAKERS, AND CAMPAIGNS 73 (Mary Crowley ed., The Hastings Center, 2008).

<sup>77</sup> LORI B. ANDREWS ET AL., GENETICS: ETHICS, LAW, AND POLICY 301 (2015); *What Are the Types of Genetic Tests?*, U.S. NAT'L LIBRARY OF MED., <https://ghr.nlm.nih.gov/primer/testing/uses> (published Feb. 26, 2019).

*What Genetic Testing Teaches About Long-Term Predictive Health  
Analytics Regulation*

Claire identified a genetic mutation,<sup>78</sup> BRCA1, that is linked to breast and ovarian cancer, as is BRCA2, which was discovered shortly thereafter.<sup>79</sup> Since then, scientists have discovered a myriad of genetic abnormalities that can increase disease vulnerabilities and have developed predictive genetic tests for some of them.<sup>80</sup> For example, predictive testing can be done for early-onset familial Alzheimer's disease, a variety of cancers, hereditary hemochromatosis (a disorder causing iron overload), Huntington's Disease, and more.<sup>81</sup>

## B. Genetic Testing Concerns

The advent of genetic testing raised numerous concerns that were vigorously debated and catapulted professional and governmental bodies into action. Academics wrote hundreds of articles about genetic testing, and law reviews dedicated entire symposium issues to the subject.<sup>82</sup> In 1995 the National Institutes

---

<sup>78</sup> A mutation "is a permanent alteration in the DNA sequence that makes up a gene, such that the sequence differs from what is found in most people." *What Is a Gene Mutation and How Do Mutations Occur?*, U.S. NAT'L LIBRARY OF MED., <https://ghr.nlm.nih.gov/primer/mutationsanddisorders/genemutation> (published Feb. 26, 2019).

<sup>79</sup> Nancy Press, *supra* note 76, at 73; Lydia Ramsey, *Over a 40-Year Career, This 'Stubborn Scientist' Helped Change the Way We Think about Cancer and Genetics*, BUS. INSIDER (Nov. 16, 2017), <https://www.businessinsider.com/mary-claire-king-and-the-impact-brca-genes-had-on-cancer-genetics-2017-11>; *BRCA1 & BRCA2 Genes: Risk for Breast & Ovarian Cancer*, MEM'L SLOAN KETTERING CANCER CTR., <https://www.mskcc.org/cancer-care/risk-assessment-screening/hereditary-genetics/genetic-counseling/brca1-brca2-genes-risk-breast-ovarian> (last visited Feb. 27, 2019).

<sup>80</sup> ANDREWS ET AL., *supra* note 77, at 301; Burke, *supra* note 74, at 1870.

<sup>81</sup> Burke *supra* note 74, at 1870; *Learning About Huntington's Disease*, NAT'L HUMAN GENOME RESEARCH INST., <https://www.genome.gov/10001215/learning-about-huntingtons-disease/> (last updated: Nov. 17, 2011); *What Are the Types of Genetic Tests?*, *supra* note 77.

<sup>82</sup> See, e.g., *The Fifth Annual Health Law Symposium, Communities of Color and Genetic Testing: Purpose, Voice, & Values*, 27 SETON HALL L. REV. 887 (1997); Symposium, *The Genetics Revolution: Conflicts, Challenges and Conundra*, 28 AM. J.L. & MED. 145 (2002); Symposium, *The Human Genome Initiative and the Impact of Genetic Testing and Screening Technologies*, 17 AM. J.L. & MED. 1 (1991); Symposium, *Is There a Pink Slip in Your Genes? Genetic Discrimination in Employment and in Health Insurance*, 16 J.L. & HEALTH 1 (2001-02); Symposium, *Legal and Ethical Issues Raised by the Human Genome Project*, 29 HOUS. L. REV. 1 (1992); Symposium, *Legal Liabilities at the Frontier of Genetic Testing* (pts. 1 & 2), 41 JURIMETRICS 1 (2000), 41 JURIMETRICS 145 (2001); Symposium, *Living in the Genetic Age: New Issues, New Challenges*, 3 ST. LOUIS U. J. HEALTH L. & POL'Y 1 (2009); *The Randolph W. Thrower Symposium, Law & Human Genetics on the Threshold of the New Millennium*, 49 EMORY L.J. 745 (2000); Symposium, *Re-defining Disability: Legal Protections for Individuals with HIV, Genetic Predispositions to Disease, or Asymptomatic Diseases*, 3 J. HEALTH CARE L. & POL'Y 225 (2000); Symposium, *Testing and Telling?*

of Health-Department of Energy Joint Working Group on the Ethical, Legal and Social Implications of Human Genome Research established a Task Force on Genetic Testing.<sup>83</sup> The American Academy of Ophthalmology convened such a task force as well.<sup>84</sup> For purposes of illustration, this section will focus on three areas of concern out of the many that were considered: clinical validity and predictive accuracy; privacy and discrimination; and psychological harms.

### ***1. Clinical Validity and Predictive Accuracy***

Experts worry about the clinical validity and predictive accuracy of genetic test results.<sup>85</sup> Many genetic tests identify only a fraction of genetic mutations that can cause a disease because researchers have yet to discover other mutations or because the price of more comprehensive testing is too high.<sup>86</sup> Moreover, although a subgroup of patients may have an inherited form of a disease such as cancer, many others will develop the disease because of environmental or other triggers without having genetic mutations.<sup>87</sup> Individuals who undergo genetic testing and receive negative results may mistakenly conclude that they are immune to the disease at issue. Thus, a woman who is found not to have the BRCA1 or BRCA2 mutation may foolishly decline appropriate screening measures, such as routine mammograms and gynecological exams. In truth, only five to ten percent of breast and ovarian cancers are hereditary, so the vast majority of these diseases have no genetic link.<sup>88</sup>

A further risk is that the opposite will occur. An individual who receives a positive genetic test result may panic and take unnecessarily aggressive preventive measures.<sup>89</sup> Many genetic mutations are not completely penetrant, that is, not all individuals with the abnormality will develop the disease at issue.<sup>90</sup> For

---

*Implications for Genetic Privacy, Family Disclosure and the Law*, 1 J. HEALTH CARE L. & POL'Y 301 (1998).

<sup>83</sup> *Task Force on Genetic Testing*, NAT'L HUMAN GENOME RESEARCH INST. (Apr. 1995), <https://www.genome.gov/10001808/genetic-testing-task-force/>.

<sup>84</sup> See Edwin M. Stone et al., *Recommendations for Genetic Testing of Inherited Eye Diseases*, 119 *Ophthalmology* 2408 (2012).

<sup>85</sup> Burke *supra* note 74, at 1871.

<sup>86</sup> *Id.*

<sup>87</sup> *BRCA1 & BRCA2 Genes: Risk for Breast & Ovarian Cancer*, *supra* note 79.

<sup>88</sup> *Id.*

<sup>89</sup> ANDREWS ET AL., *supra* note 77, at 301.

<sup>90</sup> *What Are Reduced Penetrance and Variable Expressivity?*, U.S. NAT'L LIBRARY OF MED.,

example, a woman who tests positive for the BRCA1 mutation has only a fifty-five to sixty five percent chance of developing breast cancer by the age of seventy.<sup>91</sup> Women who fully understand the meaning of their test results and the extent of their risk may or may not want to undergo prophylactic radical mastectomies, and either decision would be rational.

Physicians and patients who use genetic testing must be fully educated about how to interpret test results and the limitations of the information they reveal. It is all too easy to misconstrue test outcomes and attribute more certainty to genetic predictions than they warrant. Such misunderstandings can lead to consequential medical treatment missteps.

## **2. Privacy and Discrimination**

A dearth of regulation that protected patients against medical privacy violations and genetic discrimination led to significant concern in legal and policy circles for several decades.<sup>92</sup> Until 2003, there was no federal law that safeguarded the privacy of health information in general, let alone genetic information in particular.<sup>93</sup> Thus, federal law did not prohibit anyone who possessed genetic information from disclosing it to third parties. At the state level, a patchwork of statutes offered varying degrees of genetic privacy protections in some states.<sup>94</sup> Moreover, until the passage of the Genetic Information Nondiscrimination Act in 2008,<sup>95</sup> no federal law prohibited third parties such as employers and health insurers

---

<https://ghr.nlm.nih.gov/primer/inheritance/penetranceexpressivity> (published Feb. 26, 2019) (“Penetrance refers to the proportion of people with a particular genetic change (such as a mutation in a specific gene) who exhibit signs and symptoms of a genetic disorder.”).

<sup>91</sup> *BRCA1 and BRCA2*, SUSAN G. KOMEN, <https://ww5.komen.org/BreastCancer/BRCA1andBRCA2.html> (updated Dec. 7, 2018).

<sup>92</sup> See generally, Eric Mills Holmes, *Solving the Insurance/Genetic Fair/Unfair Discrimination Dilemma in Light of the Human Genome Project*, 85 KY. L.J. 503 (1997); Pauline T. Kim, *Genetic Discrimination, Genetic Privacy: Rethinking Employee Protections for a Brave New Workplace*, 96 NW. U. L. REV. 1497 (2002).

<sup>93</sup> HOFFMAN, *supra* note 26, at 62 (discussing the HIPAA Privacy Rule, which became effective in 2003); Karen H. Rothenberg, *Breast Cancer, The Genetic “Quick Fix,” and The Jewish Community: Ethical, Legal, and Social Challenges*, 7 HEALTH MATRIX 97, 115 (1997).

<sup>94</sup> Lori B. Andrews, *A Conceptual Framework for Genetic Policy: Comparing the Medical, Public Health, and Fundamental Rights Models*, 79 WASH. U. L.Q. 221, 280 (2001); Rothenberg, *supra* note 93, at 115-18;

<sup>95</sup> Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110-233, 122 Stat. 881 (codified as amended in scattered sections of 29 & 42 U.S.C.).



from demanding that individuals provide genetic information or from discriminating on its basis.<sup>96</sup> Some states took the lead and passed genetic discrimination legislation as early as the 1990s, but the protections they offered were inconsistent and often limited.<sup>97</sup>

Without comprehensive privacy protection, sensitive genetic information could end up in the hands of third parties that could use it to promote their own agendas. The prospect of genetic discrimination generated a plethora of literature and many heated academic and policy debates.<sup>98</sup>

For example, workers worried that employers would obtain genetic data through pre-employment or post-employment medical examinations.<sup>99</sup> Once they learned of individuals' genetic abnormalities, employers could decide to reject them, fire them, demote them, or take other adverse actions with impunity.<sup>100</sup>

Americans also were apprehensive about the impact of genetic testing on health insurance coverage.<sup>101</sup> An insurer selling an individual policy who obtained data about an applicant's disease risks could potentially decline to insure the person, raise premium prices, or dictate other adverse coverage conditions.<sup>102</sup> The same could be true for other types of insurance, such as long-term care plans.<sup>103</sup>

### **3. Psychological Harms**

A third area of concern revolved around psychological harms. Individuals who discover that they are at risk of a life-threatening disease may suffer depression and even become suicidal.<sup>104</sup> They

---

<sup>96</sup> Rothenberg, *supra* note 93, at 107-114.

<sup>97</sup> *Id.* at 108-09, 114-15.

<sup>98</sup> See *supra* notes 82-84 and accompanying text.

<sup>99</sup> Ellen R. Peirce, *The Regulation of Genetic Testing in the Workplace – A Legislative Proposal*, 46 OHIO ST. L.J. 771, 801-04 (1985).

<sup>100</sup> *Id.*

<sup>101</sup> Andrews, *supra* note 94, at 258-61.

<sup>102</sup> *Id.* at 280 (explaining that the law “does not prohibit genetic discrimination against people seeking insurance under individual plans” and “does not prohibit group insurers from charging higher rates to a whole group based on genetic information about a particular individual”); Robert Lowe, *Genetic Testing and Insurance: Apocalypse Now?* 40 DRAKE L. REV. 507, 510-11 (1991).

<sup>103</sup> Mark A. Rothstein, *Predictive Genetic Testing for Alzheimer's Disease in Long-Term Care Insurance*, 35 GA. L. REV. 707 (2001).

<sup>104</sup> Katherine A. Schneider, *Adverse Impact of Predisposition Testing on Major Life Activities: Lessons from BRCA1/2 Testing*, 3 J. HEALTH CARE L. & POL'Y 365, 369 & 372-74 (2000); Kathryn M. Kash, *Psychosocial and Ethical*

*What Genetic Testing Teaches About Long-Term Predictive Health  
Analytics Regulation*

may lose motivation to be productive in their careers, experience diminished self-esteem, and have difficulty caring for their families.<sup>105</sup> They may also decide not to get married or have children because they expect to die young and do not wish to transmit a genetic abnormality to a child.<sup>106</sup> On the other hand, some patients expect to develop inherited diseases such as breast cancer or Huntington’s disease because many of their family members suffered from the condition, and they build their lives around this assumption.<sup>107</sup> Obtaining a negative genetic test result may be just as devastating to them.<sup>108</sup> They may be confused and depressed by the need to reorient their lives and feel “survivors’ guilt” in the face of their loved ones’ suffering.<sup>109</sup>

The risk of psychological injury is particularly acute in the case of testing minors, especially for adult-onset illnesses, such as Huntington’s disease.<sup>110</sup> Experts questioned whether it was ethical to test individuals under the age of eighteen.<sup>111</sup> They also pondered who should gain access to test results and the extent to which clinicians should ask both parents and their children to consent to the testing.<sup>112</sup>

If preventive measures such as regular screening and curative medical interventions are available, genetic testing of children can be justified and beneficial.<sup>113</sup> However, in the absence of such measures, many condemn testing as potentially devastating to minors and their family members. Knowing that they live in the shadow of an impending illness could ravage minors’ psychological

---

*Implications of Defining Genetic Risk for Cancers*, ANNALS N.Y. ACAD. SCI. 41, 45-6 (1995) (discussing “psychological issues in women at genetic risk”).

<sup>105</sup> Schneider, *supra* note 104, at 369 & 372-74.

<sup>106</sup> *Id.* at 376.

<sup>107</sup> *Id.* at 374.

<sup>108</sup> *Id.*

<sup>109</sup> *Id.*

<sup>110</sup> Kimberly A. Quaid, *Genetic Testing for Huntington Disease*, in 144 HANDBOOK OF CLINICAL NEUROLOGY 113, 115-17 (A.S. Feigin & K.E. Anderson eds. 2017). For information about Huntington’s disease, see *Huntington’s disease*, MAYO CLINIC, <https://www.mayoclinic.org/diseases-conditions/huntingtons-disease/symptoms-causes/syc-20356117> (last visited Mar. 3, 2019).

<sup>111</sup> ANDREWS ET AL., *supra* note 77, at 331.

<sup>112</sup> *Id.*

<sup>113</sup> Am. Soc’y of Human Genetics Bd. of Dirs. & Am. Coll. of Med. Genetics Bd. of Dirs., *Points to Consider: Ethical, Legal, and Psychosocial Implications of Genetic Testing in Children and Adolescents*, 57 AM. J. HUMAN GENETICS 1233, 1233 (1995).

wellbeing.<sup>114</sup> Likewise, discovery of a child's genetic abnormality may upend family dynamics as the affected child is treated either as more precious than others or less favorably because the child does not have a promising future.<sup>115</sup>

### **C. Legal and Policy Interventions**

As genetic testing became increasingly common, legislators and other policy-makers implemented a variety of measures to address the concerns that it raised. This section will focus on three of these: federal and state anti-discrimination legislation, the HIPAA Privacy Rule, and self-regulation mechanisms.

#### **1. State and Federal Anti-Discrimination Legislation**

States began enacting legislation that prohibits genetic discrimination as early as the 1970s.<sup>116</sup> Early laws focused on protecting individuals with the sickle cell trait.<sup>117</sup> In 1991, Wisconsin was the first state to enact a more comprehensive statute.<sup>118</sup> Thereafter, the vast majority of states enacted genetic anti-discrimination statutes, though they varied significantly in scope and contents.<sup>119</sup> As applied to health insurers, these laws formulated restrictions related to using genetic information to determine coverage eligibility or premium levels, requiring applicants to undergo genetic testing, or disclosing genetic information to others without consent.<sup>120</sup> As applied to employers, the laws prohibited employers from discriminating on the basis of genetic information and from requesting, requiring, or obtaining genetic information.<sup>121</sup>

---

<sup>114</sup> *Id.* at 1235-36. At the same time, minors whose test results are negative “may develop ‘survivor guilt,’ based on the knowledge that one or more of their siblings will develop-and perhaps die from-a serious genetic disease.” *Id.* at 1236.

<sup>115</sup> *Id.* at 1236.

<sup>116</sup> ANDREWS ET AL., *supra* note 77, at 776.

<sup>117</sup> *Id.*

<sup>118</sup> *Id.* See Wisconsin Fair Employment Law, WIS. STAT. § 111.372 (2018).

<sup>119</sup> *Genetics and Health Insurance State Anti-Discrimination Laws*, NAT'L CONFERENCE OF STATE LEGISLATURES, <http://www.ncsl.org/research/health/genetic-nondiscrimination-in-health-insurance-laws.aspx> (last updated Jan. 2008) [hereinafter State Health Insurance Discrimination Laws]; *Genetic Employment Laws*, NAT'L CONFERENCE OF STATE LEGISLATURES, <http://www.ncsl.org/research/health/genetic-employment-laws.aspx> (last updated Jan. 2008) [hereinafter State Employment Discrimination Laws].

<sup>120</sup> State Health Insurance Discrimination Laws, *supra* note 119.

<sup>121</sup> State Employment Discrimination Laws, *supra* note 119.

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

The United States Congress considered genetic discrimination bills for thirteen long years.<sup>122</sup> Finally, President George W. Bush signed the Genetic Information Nondiscrimination Act (GINA) into law on May 21, 2008.<sup>123</sup> GINA applies to the use of predictive genetic information by health insurers and employers. The law does not cover those who already manifest symptoms of a genetic disease.<sup>124</sup>

Title I of the Act prohibits genetic discrimination in health insurance. Health insurers offering group plans may not modify premium prices and contribution amounts based on genetic information.<sup>125</sup> Insurers offering individual health plans may not require genetic testing or use genetic information to establish rules for eligibility, premium prices, or contribution amounts, or to apply preexisting condition exclusions for coverage.<sup>126</sup>

GINA's Title II focuses on employment discrimination. It establishes that employers may not discriminate against employees in hiring, firing, or other employment practices based on genetic information.<sup>127</sup> The law defines "genetic information" as including genetic testing of both individuals and their family members and family disease histories.<sup>128</sup> Furthermore, Title II prohibits employers from attempting to obtain genetic information about applicants or employees by requesting, requiring, or purchasing it.<sup>129</sup>

GINA has many critics who decry its arguably anemic protections.<sup>130</sup> For example, it applies only to health insurers and

---

<sup>122</sup> ANDREWS ET AL., *supra* note 77, at 777.

<sup>123</sup> Genetic Information Nondiscrimination Act of 2008, *supra* note 95; *President Bush Signs the Genetic Information Nondiscrimination Act of 2008*, NAT'L HUMAN GENOME RESEARCH INST. (May 21, 2008), <https://www.genome.gov/27026050/president-bush-signs-the-genetic-information-nondiscrimination-act-of-2008/>.

<sup>124</sup> Mark A. Rothstein, *GINA's Beauty Is Only Skin Deep*, GENE WATCH, Apr.-May 2009, at 9, 10 ("The problem is that GINA only applies to asymptomatic individuals.").

<sup>125</sup> 29 U.S.C. § 1182(b)(3) (2010).

<sup>126</sup> 42 U.S.C. § 300gg-53 (2010).

<sup>127</sup> 42 U.S.C. § 2000ff-1(a) (2010).

<sup>128</sup> 42 USC § 2000ff(4)(A) (2010) (defining genetic information as "(i) an individual's genetic tests, (ii) the genetic tests of an individual's family members, and (iii) the manifestation of a disease or disorder in an individual's family members")

<sup>129</sup> 42 U.S.C. § 2000ff-1(b) (2010).

<sup>130</sup> *See, e.g.*, Bradley A. Areheart & Jessica L. Roberts, *GINA, Big Data, and the Future of Employee Privacy*, 128 YALE L.J. 710, 745 (2019) (noting that "the scholarly reaction to GINA has been almost entirely negative"); Russell Korobkin

employers rather than to all parties that might possess genetic information (such as life or disability insurers) and might subject individuals to discrimination on its basis.<sup>131</sup> GINA also does not cover a range of non-genetic biologic information that may be of interest to third parties, such as epigenetic markers and the microbiome.<sup>132</sup>

A full analysis of GINA or parallel state legislation is beyond the scope of this article. My point here is merely that legislators recognized that genetic testing could yield both benefits and serious risks. They were sufficiently thoughtful and concerned about those risks to enact statutory interventions, however imperfect.

## **2. The HIPAA Privacy Rule**

The HIPAA Privacy Rule, which went into effect in 2003, is a set of federal regulations that address the privacy of health information.<sup>133</sup> The Privacy Rule establishes that, with some exceptions, covered entities<sup>134</sup> must obtain patients' permission to disclose their protected health information to third parties.<sup>135</sup> As of 2013, "health information" explicitly includes genetic information.<sup>136</sup>

Under the HIPAA Privacy Rule, covered entities must allow patients to view and obtain copies of their health records and receive

---

& Rahul Rajkumar, *The Genetic Information Nondiscrimination Act — A Half-Step Toward Risk Sharing*, 359 NEW ENG. J. MED. 335, 337 (2008) (criticizing "[t]he arbitrary nature of the categories GINA creates"); Rothstein, *supra* note 124, at 9 ("Unfortunately, the protections afforded individuals under either state laws prohibiting genetic discrimination in health insurance or GINA are not particularly robust or valuable.").

<sup>131</sup> See *supra* notes 125-129 and accompanying text.

<sup>132</sup> Areheart & Roberts, *supra* note 130, at 748-49.

<sup>133</sup> 45 C.F.R. §§ 160.101–534 (2018); *HIPAA for Professionals*, U.S. DEP'T OF HEALTH & HUMAN SERVS., <https://www.hhs.gov/hipaa/for-professionals/index.html> (last reviewed June 16, 2017). The HIPAA Privacy Rule was promulgated pursuant to the Health Insurance Portability and Accountability Act of 1996 and was amended in accordance with the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. See 42 U.S.C. §§ 1320d to 1320d-9 (2010); Health Information Technology for Economic and Clinical Health (HITECH) Act, Title XIII of Division A and Title IV of Division B of the American Recovery and Reinvestment Act of 2009 (ARRA), Pub. L. No. 111-5 (Feb. 17, 2009).

<sup>134</sup> See *infra* notes 142-143 for definition of covered entities.

<sup>135</sup> 45 C.F.R. §§ 164.508–.510 (2018).

<sup>136</sup> 45 CFR § 160.103 (2018); *Genetic Information Privacy*, ELEC. FRONTIER FOUND., <https://www EFF.ORG/ISSUES/genetic-information-privacy> (last visited Mar. 6, 2019).

*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

an accounting of disclosures of their protected health information.<sup>137</sup> In addition, patients can ask health care providers to correct errors in their medical records or to use their health data restrictively.<sup>138</sup> In addition, covered entities that suffer privacy breaches of unsecured data, such as hacking occurrences, must notify affected individuals and the Department of Health and Human Services, and in instances of large breaches, must also notify the media.<sup>139</sup>

The related HIPAA Security Rule, which became effective in 2005, promotes secure storage and processing of electronic health information (EHI).<sup>140</sup> It delineates administrative, physical, and technical safeguards to protect EHI's confidentiality, integrity, and availability.<sup>141</sup>

The HIPAA Privacy and Security Rules offer valuable protections to American patients. However, like the genetic anti-discrimination statutes, they are limited in scope and have been attacked by critics. For example, "covered entities" include only health plans, health care clearinghouses, health care providers who transmit health information electronically for purposes of HIPAA-relevant transactions, and their business associates.<sup>142</sup> Therefore, other parties that possess and handle health information, such as data brokers and marketers, need not comply with the Rules' privacy and security mandates.<sup>143</sup>

Another noteworthy regulatory gap in the HIPAA Privacy and Security Rules is the absence of a private cause of action.<sup>144</sup> Thus, individuals whose health data is breached cannot sue wrongdoers for damages under federal law no matter what consequences they suffer. Instead, the regulations leave enforcement solely in the hands of the Department of Health and Human Services Office for Civil Rights (OCR) and state attorneys

---

<sup>137</sup> 45 C.F.R. §§ 164.520(b)(1)(iv), 164.528 (2018).

<sup>138</sup> 45 C.F.R. §§ 164.520(b)(1)(iv), 164.522 (2018).

<sup>139</sup> 45 C.F.R. §§ 164.400–.408 (2018). Entities must notify the media if a breach involves "more than 500 residents of a State or jurisdiction." *Id.* § 164.408. Unsecured protected health information means information "that is not rendered unusable, unreadable, or indecipherable to unauthorized persons through the use of a [specified] technology or methodology," such as encryption.

<sup>140</sup> Sharon Hoffman & Andy Podgurski, *In Sickness, Health, and Cyberspace: Protecting the Security of Electronic Private Health Information*, 48 B.C. L. REV. 331, 335-36 (2007).

<sup>141</sup> 45 C.F.R. §§ 164.302–.318 (2018).

<sup>142</sup> 45 C.F.R. §§ 160.102–.103 (2018); 42 U.S.C. § 17934 (2010).

<sup>143</sup> See HOFFMAN, *supra* note 26, at 73.

<sup>144</sup> See HOFFMAN, *supra* note 26, at 75.

general offices, which may or may not have adequate staffing and resources for robust prosecutorial activities.<sup>145</sup>

In addition, the HIPAA Privacy Rule features numerous exceptions. Covered entities can disclose patients' medical data for purposes of treatment, payment, and health care operations without patient authorization.<sup>146</sup> Thus, the regulations permit physicians to consult colleagues about patients and to ask administrators to review records for billing or other office-related purposes without the patients' knowledge. The Rule exempts additional disclosures as well, such as those made for law enforcement, public health, and other listed purposes.<sup>147</sup> In general, these exceptions are reasonable and sound. However, patients should understand that they are often ignorant of who is viewing their health data and for what purpose.

The HIPAA Privacy and Security Rules fall short of providing American patients with comprehensive protection. Nevertheless, they constitute important advances in the privacy arena and address some of the concerns raised by genetic testing.

### **3. Self-Regulation**

Genetic testing professionals also engage in self-regulation, formulating practice guidelines and deferring testing until they have educated patients about its potential consequences. For example, The American Society of Breast Surgeons issued a "Consensus Guideline on Genetic Testing for Hereditary Breast Cancer."<sup>148</sup> The guidance formulated recommendations for genetic testing and discussed testing limitations.<sup>149</sup> The American Society of Human Genetics Board of Directors and the American College of Medical Genetics Board of Directors published a document entitled "Points to Consider: Ethical, Legal, and Psychosocial Implications of Genetic Testing in Children and Adolescents."<sup>150</sup> The guidance recommended against testing children for adult-onset diseases if they will derive no medical or psychological benefit from being

---

<sup>145</sup> 45 C.F.R. § 160.306 (2018); 42 U.S.C. § 1320d-5(d) (2010).

<sup>146</sup> 45 C.F.R. § 164.506 (2018).

<sup>147</sup> 45 C.F.R. §§ 164.502, .512 (2018).

<sup>148</sup> *Consensus Guideline on Genetic Testing for Hereditary Breast Cancer*, AM. SOC'Y OF BREAST SURGEONS, [https://www.breastsurgeons.org/about/statements/PDF\\_Statements/Hereditary\\_Genetic\\_Testing\\_Patients\\_With\\_Without\\_Breast\\_Cancer.pdf](https://www.breastsurgeons.org/about/statements/PDF_Statements/Hereditary_Genetic_Testing_Patients_With_Without_Breast_Cancer.pdf) (last visited Mar. 7, 2019).

<sup>149</sup> *Id.*

<sup>150</sup> Am. Soc'y of Human Genetics Bd. of Dirs. & Am. Coll. of Med. Genetics Bd. of Dirs., *supra* note 113.

tested as minors.<sup>151</sup> The American College of Medical Genetics and Genomics has issued practice guidelines related to genetic testing for numerous conditions.<sup>152</sup>

Clinicians typically offer genetic counseling to patients who are considering genetic testing.<sup>153</sup> Counseling is designed to ensure that patients make fully informed decisions about pursuing testing, in light of the benefits and risks that exist in their particular circumstances.<sup>154</sup> A variety of health care providers can educate patients about genetic testing, but a growing number of practices include professional genetic counselors with master's degrees.<sup>155</sup> The American Board of Genetic Counselors has certified over 4,000 genetic counselors thus far.<sup>156</sup> Thus, rather than rushing to test patients after only a brief discussion, responsible clinicians exercise a degree of self-restraint and take steps to ensure that patients provide meaningful and genuinely informed consent to the procedure.

#### **IV. LONG-TERM PREDICTIVE HEALTH ANALYTICS CONCERNS**

Like genetic testing, long-term predictive health analytics are fraught with risks, but these are garnering too little attention.<sup>157</sup> This section highlights three areas of concern, though this list is far from comprehensive. It focuses on psychological harms, privacy and discrimination, and erroneous predictions.

##### **A. Psychological Harms**

---

<sup>151</sup> *Id.* at 1233. See also, Quaid, *supra* note 110, at 115-17 (discussing testing guidelines for Huntington's disease, including those addressing predictive testing of minors).

<sup>152</sup> *Practice Guidelines*, AM. COLL. OF MED. GENETICS & GENOMICS, <https://www.acmg.net/ACMG/Medical-Genetics-Practice-Resources/Practice-Guidelines.aspx> (last visited Mar. 7, 2019).

<sup>153</sup> Burke, *supra* note 74, at 1873.

<sup>154</sup> *Id.*

<sup>155</sup> ANDREWS ET AL., *supra* note 77, at 329; *About NSGC: The Leading Voice for Genetic Counselors*, NAT'L SOC'Y OF GENETIC COUNSELORS, <https://www.nsgc.org/page/about-nsgc> (last visited Mar. 7, 2019).

<sup>156</sup> *Mission Statement, Purpose and Values*, AM. BD. OF GENETIC COUNSELORS, <https://www.abgc.net/about-abgc/mission-history/> (last visited Mar. 7, 2019).

<sup>157</sup> I. Glenn Cohen & Harry S. Graver, *Cops, Docs, and Code: A Dialogue between Big Data in Health Care and Predictive Policing*, 51 U.C. DAVIS L. REV. 437, 446 (2017) ("The legal literature on predictive analytics in health care is at this moment less robust than that on predictive policing, although that is changing.").



*What Genetic Testing Teaches About Long-Term Predictive Health Analytics Regulation*

Predictions of future ailments based on predictive health analytics can be just as traumatizing as predictions based on genetic testing.<sup>158</sup> Individuals who learn from their doctors that they are likely to develop heart disease, dementia, or psychosis in the future might feel that the news is devastating.<sup>159</sup> As a result, they could have difficulty concentrating on work, experience strain in their relationships, or even become clinically depressed or suicidal.<sup>160</sup> Like genetic testing of children, predictive health analytics involving minors raise particularly troubling questions.<sup>161</sup> Worrisome predictions can adversely impact children's futures and disrupt family dynamics.<sup>162</sup>

Physicians who identify certain individuals as vulnerable to opioid addiction, cognitive decline, or suicide<sup>163</sup> may treat those patients differently, to the patients' detriment. For example, they may refuse to provide potential opioid addicts with needed pain medication.<sup>164</sup> They may also relate to patients at risk of dementia or suicide poorly, treating them as cognitively compromised or lacking autonomy even when they are fully competent.<sup>165</sup> So too they may try to drive patients who are labelled as potentially high-risk and high-cost away from their practices.<sup>166</sup>

It is also likely that many individuals will obtain distressing health predictions not from their doctors, but from commercial enterprises, without being aware in advance that anyone has assessed their health risks. Data brokers sell health information to interested buyers, and companies such as LexisNexis and Acxiom have already begun to engage in predictive health analytics.<sup>167</sup> Marketers will likely be eager to obtain health predictions about patients in order to tailor their marketing materials effectively.<sup>168</sup> Imagine individuals receiving the news that they are at risk of cognitive decline through an electronic advertisement urging them to purchase memory-enhancing products! People who do not have

---

<sup>158</sup> See *supra* Part III.B.3.

<sup>159</sup> See *supra* Part II.B.

<sup>160</sup> See *supra* notes 104-106 and accompanying text.

<sup>161</sup> See *supra* notes 111-115 and accompanying text.

<sup>162</sup> *Id.*

<sup>163</sup> See *supra* Part II.B.

<sup>164</sup> Ravindranath, *supra* note 66.

<sup>165</sup> Marks, *supra* note 12, at \_\_\_ (“Patients with mental illnesses often report feeling dehumanized and dismissed by healthcare providers.”).

<sup>166</sup> Cohen et al., *supra* note 5, at 1141.

<sup>167</sup> See *supra* notes 66-73 and accompanying text.

<sup>168</sup> Cobb, *supra* note 73; HOFFMAN, *supra* note 26, at 60.

the support of a physician and do not receive a clear, medical explanation of the prediction and its degree of certainty will be all the more vulnerable to distress and misunderstandings.

## **B. Privacy and Discrimination**

Because the HIPAA Privacy Rule governs only a subset of parties that possess health information,<sup>169</sup> not all predictive health analytics outcomes will be subject to privacy protections.<sup>170</sup> Entities that are not health plans, health care clearinghouses and health care providers or their business associates are not legally bound to refrain from disclosing health information about patients.<sup>171</sup> Thus, data brokers are permitted to sell health-related information to marketers.<sup>172</sup> Moreover, entities that are not covered by HIPAA could disclose and publicize individually identifiable predictive health analytics results. One can imagine the media obtaining predictions about entertainers and politicians that can cause significant embarrassment and even ruin careers. Predictions about ordinary people could likewise be widely publicized through social media and be available to anyone with an electronic device.

It is also noteworthy that non-covered entities are not subject to the requirements of the HIPAA Security Rule.<sup>173</sup> Thus, such entities may be tempted to use security shortcuts and may be vulnerable to hacking and other data breaches.<sup>174</sup> Data stored by commercial enterprises for predictive health analytics purposes, consequently, may be more vulnerable to privacy violations than HIPAA-protected health information.

Given the HIPAA Privacy Rule's limitations, individuals' health predictions can easily land in the hands of third parties who may use them to further their own economic agendas.<sup>175</sup> Employers, lenders,

---

<sup>169</sup> See *supra* note 142 and accompanying text.

<sup>170</sup> Marks *supra* note 12, at \_\_\_ (discussing privacy risks related to suicide predictions).

<sup>171</sup> *Id.*; 45 C.F.R. §§ 160.102–103 (2018); 42 U.S.C. § 17934 (2010).

<sup>172</sup> Erin McCann, *What HIPAA Doesn't Cover*, HEALTHCARE IT NEWS (July 16, 2014), <https://www.healthcareitnews.com/news/what-hipaa-doesnt-cover>.

<sup>173</sup> Sharona Hoffman & Andy Podgurski, *supra* note 140, at 344-47.

<sup>174</sup> *Id.* at 332-34 (discussing various data breaches); Topol, *supra* note 4, at 52 (noting “the risk of deliberate hacking of an algorithm to harm people at a large scale”).

<sup>175</sup> HOFFMAN, *supra* note 26, at 59-60 (listing a variety of parties that could be interested in people's health data).

life insurers, and others with a stake in individuals' future wellbeing may be interested in predictions about individuals' future health status.<sup>176</sup> Employers for example, are interested in employees who will not have productivity or absenteeism problems and will not generate high health insurance costs.<sup>177</sup> They may be very tempted to reject or terminate workers who they believe to be at high risk of becoming seriously ill in the coming years. Similarly, lenders seek borrowers who will remain able to work and pay off their loans, and life insurers may use predictive information about applicants to make eligibility or pricing decisions.<sup>178</sup>

Currently, the anti-discrimination laws do not prohibit employers and others from discriminating based on predictions of future health problems (other than predictions based on genetic information, which are covered by GINA).<sup>179</sup> The Americans with Disabilities Act (ADA), the primary federal disability discrimination law, prohibits discrimination related only to:

- (A) a physical or mental impairment that substantially limits one or more major life activities of such individual;
- (B) a record of such an impairment; or
- (C) being regarded as having such an impairment.<sup>180</sup>

Consequently, it does not reach discrimination based on *future* physical or mental impairments or disabilities. This legislative gap creates worrisome opportunities for discrimination based on disease predictions.

Predictive health analytics may also perpetuate other types of discrimination, such as sex or race-based discrimination. Amazon's effort to develop artificial intelligence-driven software to identify the best job candidates illustrates the point.<sup>181</sup> Because the predictive model's training data were past resumes submitted to Amazon mostly by men, the program was biased against women and

---

<sup>176</sup> Sharon Hoffman, *Big Data's New Discrimination Threats: Amending the Americans with Disabilities Act to Cover Discrimination Based on Data-Driven Predictions of Future Disease*, in *BIG DATA, HEALTH LAW, AND BIOETHICS* 85, 85 (I. Glenn Cohen et al. eds., 2018).

<sup>177</sup> *Id.* at 86.

<sup>178</sup> HOFFMAN, *supra* note 26, at 60; Marks, *supra* note 12, at \_\_\_.

<sup>179</sup> Hoffman *supra* note 176, at 92–94.

<sup>180</sup> 42 U.S.C. § 12102(1) (2010).

<sup>181</sup> Katherine Maher, *Without Humans, A.I. Can Wreak Havoc*, N.Y. TIMES, Mar. 12, 2019, <https://www.nytimes.com/2019/03/12/opinion/artificial-intelligence-wikipedia.html?smid=nytcore-ios-share>.

concluded men were preferable job candidates.<sup>182</sup> It is just as likely that predictive models in the healthcare arena will be biased and wrongly conclude that women are at higher risk of various health problems. Likewise, if companies such as LexisNexis and Acxiom base predictive models on variables that include criminal records and voter registration data,<sup>183</sup> they could disproportionately identify certain minorities as high-risk patients.

### **C. Erroneous Predictions**

To make matters worse, the results of predictive health analytics can often be wrong.<sup>184</sup> In one illustrative example outside the health field, scientists produced “fooling images” that were completely unrecognizable to humans, but deep neural networks (a form of machine learning) believed “with near certainty” that they were familiar objects.<sup>185</sup> Flawed outcomes can stem from a variety of problems.

One reason can be human error in the algorithm or its implementation. Big data used to train computers and develop learning algorithms can be rife with inaccuracies and data gaps or otherwise be a poor fit for the task at hand.<sup>186</sup> Poor data quality will inevitably lead to poor data-driven artificial intelligence algorithms, consistent with the “garbage in-garbage out” principle.<sup>187</sup>

Even with a correct learning algorithm, the performance the predictive model exhibits using the training data<sup>188</sup> may not

---

<sup>182</sup> *Id.* See *infra* Part IV.C for a discussion of bias.

<sup>183</sup> Ravindranath, *supra* note 66.

<sup>184</sup> Ian A. Scott, *Hope, Hype and Harms of Big Data*, 49 INTERNAL MED. J. 126, 127 (2019).

<sup>185</sup> Anh Nguyen et al., *Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images*, COMPUTER VISION & PATTERN RECOGNITION, IEEE, 1, 1 (2015), available at <https://arxiv.org/pdf/1412.1897.pdf>.

<sup>186</sup> Scott, *supra* note 184, at 127 (discussing numerous potential shortcomings of big data); Nilay D. Shah et al., *Big Data and Predictive Analytics: Recalibrating Expectations*, 320 JAMA 27, 28 (2018); Topol, *supra* note 4, at 51.

<sup>187</sup> P. Elliott Miller et al., *Predictive Abilities of Machine Learning Techniques May Be Limited by Dataset Characteristics: Insights from the UNOS Database*, J. CARDIAC FAILURE (forthcoming 2019) (manuscript at 4) (“Our results raise the notion that large clinical datasets might lack the accuracy and granularity needed for machine learning methodologies to uncover unique associations.”); Kun-Hsing Yu & Isaac S. Kohane, *Framing the Challenges of Artificial Intelligence in Medicine*, 28 BMJ QUAL. & SAF. 238, 239 (2019).

<sup>188</sup> See *supra* notes 28-31 and accompanying text for explanation of machine learning.

generalize to real world data because of a phenomenon called “overfitting.”<sup>189</sup> A particular model can produce accurate predictions on a set of training data but fail to provide sound predictions when deployed on new data, especially if the model is complex and the training data set was small.<sup>190</sup> Because of the dearth of training data and the large number of parameters used to construct the model, “the learned parameters are spuriously inferred” and are unlikely to generalize well to unseen data.<sup>191</sup> In other words, overfitting occurs when a predictive model fits the training data “too well.”<sup>192</sup>

Big data can also be subject to selection bias. If the data used to train learning algorithms or statistical models comes from a health system that serves particular populations (e.g. disproportionately wealthy or disadvantaged individuals) but not others, the algorithm or model may not be generalizable to all patients.<sup>193</sup> Several scholars have noted the following:

Big data has not captured certain marginalized demographics. Particularly concerning are racial minorities, people with low socioeconomic status, and immigrants. Many of the people missing from the data that come from sources such as Internet history, social media presence, and credit-card use are also missing from other sources of Big Data, such as electronic health records (EHRs) and genomic databases. The factors responsible for these gaps are diverse and include lack of insurance and the inability to access healthcare, to name just two....<sup>194</sup>

Unfortunately, it is often impossible to discern whether a predictive model is sound. Learning algorithms are often opaque because they rely on extremely complex rules and even their programmers are uncertain about how they ultimately work.<sup>195</sup>

---

<sup>189</sup> CHARU C. AGRAWAL, NEURAL NETWORKS AND DEEP LEARNING 25 (2018).

<sup>190</sup> *Id.*

<sup>191</sup> *Id.*

<sup>192</sup> SHALEV-SHWARTZ & BEN DAVID, *supra* note 21 at 16.

<sup>193</sup> Craig Konnoth, *Health Information Equity*, 165 U. PA. L. REV. 1317, 1361 (2017) (asserting that “relying on data that is biased towards certain social groups can have problematic effects”).

<sup>194</sup> Sarah E. Malanga et al., *Who's Left Out of Big Data? How Big Data Collection, Analysis, and Use Neglects Populations Most in Need of Medical and Public Health Research and Interventions*, in BIG DATA, HEALTH LAW, AND BIOETHICS 98, 98–99 (I. Glenn Cohen et al. eds., 2018).

<sup>195</sup> Price, *supra* note 19, at 430; Tokio Matsuzaki, *Ethical Issues of Artificial Intelligence in Medicine*, 55 CAL. W. L. REV. 255, 269 (2018) (“One concern is that AI [artificial intelligence] decision-making ... often has no transparency. This

Some commentators use the term “black box medicine” to describe reliance on nontransparent learning algorithms.<sup>196</sup>

Use of the terms “artificial intelligence” and “machine learning” can over-awe people. But as one commentator notes, “the only sure prediction about the future of big data and predictive analytics is that it is unlikely to live up to some of the hype.”<sup>197</sup> Therefore, algorithms, in many cases, will falsely predict that individuals will suffer particular conditions in the future, and the affected data subjects will be left to suffer the consequences.

## V. RECOMMENDATIONS

At its core, this Article is a call to action. The policy and scientific communities must not ignore the potential risks of predictive health analytics. Just as the growth of genetic testing elicited robust academic and policy debates, so too should the burgeoning phenomenon of predictive health analytics. Effective legal and policy interventions are needed to safeguard the rights of data subjects. This section recommends changes to the HIPAA Privacy and Security Rules and to the Americans with Disabilities Act. It also advocates for the implementation of other oversight and self-regulation mechanisms.

### A. Legal Interventions

Legislators should modify the laws that establish privacy and anti-discrimination mandates. This section focuses on the federal HIPAA Privacy and Security Rules and Americans with Disabilities Act, though states could make similar changes to parallel state laws.<sup>198</sup> It also briefly examines whether the Food and Drug Administration (FDA) could regulate learning algorithms used for long-term predictive health analytics.

---

means that doctors and patients are not able to know how the AI system reached the decision.”).

<sup>196</sup> Price, *supra* note 19, at 429; Topol, *supra* note 4, at 51.

<sup>197</sup> Shah et al., *supra* note 186, at 28.

<sup>198</sup> See Sharona Hoffman, *Medical Privacy and Security*, in THE OXFORD HANDBOOK OF U.S. HEALTHCARE LAW 267, 273–79 (I. Glenn Cohen et al. eds., 2015) (discussing state privacy laws); *Disability Discrimination: State Law, WORKPLACE FAIRNESS*, <https://www.workplacefairness.org/disability-discrimination-state-law> (last visited Mar. 15, 2019).

### **1. The HIPAA Privacy and Security Rules**

As I have argued in other work,<sup>199</sup> Congress and the Department of Health and Human Services should expand HIPAA's definition of "covered entity."<sup>200</sup> The need for change has become all the more urgent in light of the growing use of predictive health analytics. The federal law and regulations should adopt the language of a much broader Texas privacy statute:

"Covered entity" means any person who:

(A) for commercial, financial, or professional gain, monetary fees, or dues, or on a cooperative, nonprofit, or pro bono basis, engages, in whole or in part, and with real or constructive knowledge, in the practice of assembling, collecting, analyzing, using, evaluating, storing, or transmitting protected health information. The term includes a business associate, health care payer, governmental unit, information or computer management entity, school, health researcher, health care facility, clinic, health care provider, or person who maintains an Internet site;

(B) comes into possession of protected health information;

(C) obtains or stores protected health information under this chapter; or

(D) is an employee, agent, or contractor of a person described by Paragraph (A), (B), or (C) insofar as the employee, agent, or contractor creates, receives, obtains, maintains, uses, or transmits protected health information.<sup>201</sup>

This change would require a parallel modification to the definition of "health information."<sup>202</sup> "Health Information" should

---

<sup>199</sup> HOFFMAN, *supra* note 26, at 74.

<sup>200</sup> Hoffman & Podgurski, *supra* note 140, at 360-363.

<sup>201</sup> TEX. HEALTH & SAFETY CODE ANN. § 181.001(b)(2) (West 2017).

<sup>202</sup> "Health information" currently means:

any information, whether oral or recorded in any form or medium, that –  
(A) is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse; and

(B) relates to the past, present, or future physical or mental health or condition of an individual, the provision of health care to an individual,

*What Genetic Testing Teaches About Long-Term Predictive Health  
Analytics Regulation*

be expanded to mean “any information, recorded in any form or medium, that relates to the past, present, or future physical or mental health or condition of an individual, including health predictions, the provision of healthcare to an individual, or the past, present, or future payment for the provision of healthcare to an individual.”<sup>203</sup>

Expanding the definitions of “covered entities” and “health information” would not prevent healthcare providers from contracting with business associates such as LexisNexis or Axcion to conduct predictive health analytics so long as they did so for purposes of treatment, payment, or health care operations.<sup>204</sup> It also would not prevent data brokers from accessing much of the data they use, such as Facebook posts, shopper loyalty program records, or voter registration data.<sup>205</sup>

Nevertheless, the change would provide patients with several important benefits. First, it would prevent the newly covered entities from disclosing health predictions to third parties without the data subject’s consent. These predictions would constitute health information about individuals’ future physical or mental health conditions, and thus data brokers and other commercial enterprises could not sell them without permission to marketers, employers, and other interested parties for financial gain. Second, upon request, the newly covered entities would be bound to inform data subjects of all disclosures of their protected health information that were made.<sup>206</sup> Third, the newly covered entities would have to comply with the security mandates of the HIPAA Security Rule.<sup>207</sup> They therefore would be prohibited from storing health information and health predictions about individuals with sloppy or minimal security measures that do not adequately deter hacking.

In addition, the HIPAA Privacy and Security Rules should include a private cause of action.<sup>208</sup> Because of budgetary

---

or the past, present, or future payment for the provision of health care to an individual.

45 C.F.R. § 160.103 (2018); 42 U.S.C. § 1320d(4) (2000).

<sup>203</sup> See HOFFMAN, *supra* note 26, at 75.

<sup>204</sup> See *supra* note 146 and accompanying text.

<sup>205</sup> See *supra* Parts II.C. & D; proposed definition of “health information,” *supra* note 203 and accompanying text.

<sup>206</sup> 45 C.F.R. §§ 164.520(b)(1)(iv), .528 (2018).

<sup>207</sup> 45 C.F.R. §§ 164.302–.318 (2018); see *supra* notes 140-141 and accompanying text.

<sup>208</sup> See HOFFMAN, *supra* note 26, at 78-9; *supra* notes 143-144 and accompanying text (discussing the absence of a private cause of action).



constraints, government enforcement is often anemic.<sup>209</sup> Furthermore, it does not provide aggrieved parties with monetary relief if they have suffered injury resulting from a privacy breach.<sup>210</sup> The proposed HIPAA changes could meaningfully enhance data subjects' privacy protections and rights.

## **2. *The Americans with Disabilities Act***

I have also previously argued for a broadening of the Americans with Disabilities Act's (ADA) definition of "disability," and I renew my call for this change here.<sup>211</sup> The ADA's "regarded as" provision protects only individuals who are "being regarded as [currently] having ... an impairment."<sup>212</sup> Congress should revise the "regarded as" provision of the ADA to include individuals who "are perceived as likely to develop physical or mental impairments in the future."<sup>213</sup>

This change would prohibit employers and other parties from discriminating against individuals because of disease predictions.<sup>214</sup> It follows logically from GINA, which forbids discrimination based on a specific type of predictive data - genetic information.<sup>215</sup> In the era of predictive health analytics, there is no justification for retaining a discrepancy between GINA and the ADA. This is particularly true because predictive models can forecast inherited diseases such as heart conditions and some forms of Alzheimer's disease.<sup>216</sup> Presumably, GINA would not cover such predictions because they are not based on genetic tests or family histories.<sup>217</sup> The law should not leave individuals who are identified as being at risk for future diseases vulnerable to discrimination and protect them only if the prediction is rooted directly in genetic information. With respect to anti-discrimination mandates, genetic exceptionalism<sup>218</sup> no longer makes sense.

---

<sup>209</sup> Morgan Leigh Tendam, *The HIPAA-Pota-Mess: How HIPAA's Weak Enforcement Standards Have Led States To Create Confusing Medical Privacy Remedies*, 79 OHIO ST. L.J. 411, 421-22 (2018).

<sup>210</sup> See HOFFMAN, *supra* note 26, at 75-76.

<sup>211</sup> Sharona Hoffman, *Big Data and the Americans with Disabilities Act*, 68 HASTINGS L.J. 777, 787-88 (2017).

<sup>212</sup> 42 U.S.C. § 12102(1)(C) (2010).

<sup>213</sup> Hoffman, *supra* note 211 at 787.

<sup>214</sup> Hoffman, *supra* note 176, at 94-96.

<sup>215</sup> See *supra* Part III.C.1.

<sup>216</sup> See *supra* Parts II.B & C and note 81 and accompanying text.

<sup>217</sup> See *supra* Part III.C.1.

<sup>218</sup> Genetic exceptionalism is the belief that genetic information is special and should be treated differently from other health data. James P. Evans & Wylie

### **3. FDA Regulation?**

Genetic tests are subject to regulation by the FDA and the Centers for Medicare and Medicaid Services (CMS), which oversees clinical laboratories.<sup>219</sup> CMS would not regulate learning algorithms because no clinical laboratories are involved, and a real question exists as to whether the FDA will routinely do so.

The FDA regulates medical devices, which are defined as any “instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article which is intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals.”<sup>220</sup> Arguably, the FDA has authority to regulate learning algorithms as contrivances that are used for the diagnosis and prevention of disease.<sup>221</sup> In early 2018 the FDA in fact approved an algorithm.<sup>222</sup> It provided premarket clearance for the WAVE Clinical Platform, an early-warning system for hospitals that uses vital sign data to identify patients at risk of becoming unstable.<sup>223</sup>

Even under this liberal interpretation, however, the agency is empowered to regulate only algorithms used for medical care.<sup>224</sup> It would not have jurisdiction over predictive health analytics conducted by marketers, employers, or other parties for nonmedical purposes.<sup>225</sup>

Moreover, the FDA has traditionally refrained from regulating the practice of medicine.<sup>226</sup> Thus, it may hesitate to regulate learning algorithms when their use seems akin to medical practice.<sup>227</sup> While WAVE may be classified as a device designed to predict imminent medical crises,<sup>228</sup> long-term predictive analytics are a poorer match.

---

Burke, *Genetic Exceptionalism. Too Much of a Good Thing?*, 10 GENETICS IN MED. 500, 500 (2008).

<sup>219</sup> *Regulation of Genetic Tests*, NAT’L HUMAN GENOME RESEARCH INST., <https://www.genome.gov/10002335/regulation-of-genetic-tests/> (last updated Jan. 17, 2018).

<sup>220</sup> 21 U.S.C. § 321(h) (2012).

<sup>221</sup> Price, *supra* note 19, at 439.

<sup>222</sup> Ravi B. Parikh et al., *Regulation of Predictive Analytics in Medicine*, 363 SCIENCE 810, 811 (2019).

<sup>223</sup> *Id.*; *How WAVE Works*, EXCEL MED., <https://www.excel-medical.com/how-wave-works> (last visited Mar. 16, 2019).

<sup>224</sup> *See supra*, Parts II.B, C, & D.

<sup>225</sup> *See supra* notes 71-73 and accompanying text.

<sup>226</sup> Price, *supra* note 19, at 441.

<sup>227</sup> *Id.* at 441-42. *See also* Cortez, *supra* note 27, at 78.

<sup>228</sup> Parikh et al., *supra* note 222, at 811.

At best, they merely assist physicians in making decisions about a patient's future care.

Professor Nathan Cortez argues that predictive health analytics do not fit comfortably into any of the familiar categories of medical products, medical practice, or medical information for regulatory purposes.<sup>229</sup> He and others argue for a new regulatory paradigm.<sup>230</sup>

Resolving the controversy regarding FDA regulation of learning algorithms is beyond the scope of this paper. Suffice it to say that no governing authority regularly scrutinizes learning algorithms before they are deployed in order to ensure their quality. Long-term health predictions can significantly impact people's lives. Much more must be done to construct thoughtful, responsible legal oversight mechanisms for predictive health analytics.

## **B. Other Oversight Mechanisms**

Academics and other experts have begun building a literature about the legal and ethical implications of predictive health analytics only in recent years.<sup>231</sup> It is a long way from reaching the proportions of the genetic testing literature. Moreover, existing literature has paid little attention to long-term predictive health analytics, which raise unique concerns about psychological harms and discrimination.<sup>232</sup> Legal and bioethics scholars should no more ignore these risks than they did the risks of genetic testing. What follows is a brief discussion of potential oversight improvements for the predictive health analytics industry and medical professionals.

### ***1. Predictive Health Analytics Oversight***

A few papers have undertaken the development of initial guidelines for predictive health analytics.<sup>233</sup> For example, a panel of

---

<sup>229</sup> Cortez, *supra* note 27, at 81. *See also*, Parikh et al., *supra* note 222, at 811 (stating that "existing FDA standards do not neatly translate to advanced predictive algorithms").

<sup>230</sup> Cortez, *supra* note 27, at 81-85; Price, *supra* note 19, at 457-59 (calling for reform and suggesting that the FDA should regulate predictive health analytics in collaboration with other health care stakeholders).

<sup>231</sup> Cohen & Graver, *supra* note 157, at 446.

<sup>232</sup> *See supra* Parts IV A & B.

<sup>233</sup> *See* Ruben Amarasingham et al., *Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges*, 4 EGEMS Iss. 1, Art. 3 (2016), available at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4837887/pdf/egems1163.pdf>;

*What Genetic Testing Teaches About Long-Term Predictive Health  
Analytics Regulation*

seventeen experts proposed the following guiding principles in 2016:

1. Data Barriers: Establish mechanisms within the scientific community to support data sharing for predictive model development and testing.
2. Transparency: Set standards around e-HPA validation based on principles of scientific transparency and reproducibility.
3. Ethics: Develop both individual-centered and society-centered risk-benefit approaches to evaluate e-HPA.
4. Regulation and Certification: Construct a self-regulation and certification framework within e-HPA.
5. Education and Training: Make significant changes to medical, nursing, and paraprofessional curricula by including training for understanding, evaluating, and utilizing predictive models.<sup>234</sup>

The scholars that have pondered predictive health analytics all agree that transparency and oversight are of critical importance.<sup>235</sup> They recommend the establishment of industry-wide validation and certification mechanisms implemented by the Joint Commission, certifiers overseen by the FDA, independent institutional review boards, or other third parties.<sup>236</sup>

Experts have developed a variety of techniques to assess learning algorithms and predictive models.<sup>237</sup> A popular method for estimating prediction error is cross-validation.<sup>238</sup> Another method to assess statistical accuracy is the bootstrap method.<sup>239</sup> Most importantly, researchers should validate learning algorithms in the field, using real patients under the same conditions as those intended for the algorithm's post-approval use.<sup>240</sup> Such validation should ensure that the algorithm's predictive capability generalizes to the true target population.<sup>241</sup> Oversight bodies should consist of

---

Cohen et al., *supra* note 5; Parikh et al., *supra* note 222 (proposing five criteria “for evaluation and regulation of predictive algorithms”).

<sup>234</sup> Amarasingham et al., *supra* note 233, at 2, 9. E-HPA is electronic health predictive analytics.

<sup>235</sup> *Id.* at 5-9; Cohen et al., *supra* note 5, at 1142-43; Marks, *supra* note 12, at \_\_.

<sup>236</sup> See *supra* note 235; Cortez, *supra* note 27, at 84-85.

<sup>237</sup> TREVOR HASTIE ET AL., *THE ELEMENTS OF STATISTICAL LEARNING: DATA MINING, INFERENCE, AND PREDICTION* 219-57 (2d ed. 2009) (discussing model assessment and selection).

<sup>238</sup> *Id.* at 241-49.

<sup>239</sup> *Id.* at 249-54.

<sup>240</sup> Marks, *supra* note 12, at \_\_ (discussing suicide prediction research).

<sup>241</sup> See *supra* Part IV.C (discussing erroneous outcomes).

predictive health analytics and validation experts who can be trusted to scrutinize proposed assessment methods and ensure that they are appropriate.<sup>242</sup>

The recommendations offered thus far are sound, and experts should continue to develop and augment them in order to furnish policy-makers with proposals that are as detailed and evidence-based as possible. Oversight and quality control could prevent many erroneous predictions and save clinicians and patients considerable angst.

## ***2. Self-Regulation by the Medical Profession***

Health care professionals would be wise to adopt their own safeguards in order to minimize the hazards of long-term disease predictions for patients. To that end, physicians should receive training concerning long-term predictive health analytics so that they understand the extent to which it can be limited and uncertain. They should also counsel and educate patients before disclosing troubling health predictions to them. A process akin to genetic counseling would be very useful.<sup>243</sup> Patients should understand the advantages and disadvantages of learning about their disease risks and be able to make informed decisions about their choice.

In addition, professional organizations should develop practice guidelines regarding when it is appropriate to employ predictive health analytics and the extent to which clinicians should rely upon them.<sup>244</sup> For example, practice guidelines might recommend that clinicians refrain from obtaining certain types of predictions about children.<sup>245</sup> They might also suggest which interventions should and should not be implemented in response to predictions of suicidal ideation, clinical depression, opioid addiction or other ailments. Education, counseling, and practice guidelines could go far in maximizing the benefits and minimizing the risks of predictive health analytics.

---

<sup>242</sup> Ewout W. Steyerberga & Frank E. Harrell Jr, *Prediction Models Need Appropriate Internal, Internal-External, and External Validation*, 69 J. CLINICAL EPIDEMIOLOGY 245, 245-247 (2016).

<sup>243</sup> See *supra* notes 153-156 and accompanying text (discussing genetic counseling).

<sup>244</sup> See *supra* notes 148-152 and accompanying text (discussing practice guidelines for genetic testing).

<sup>245</sup> See *supra* notes 113-115 and accompanying text (discussing genetic testing of children).

#### **IV. CONCLUSION**

We should not be blinded by enthusiasm for long-term predictive health analytics or be naively seduced by technologies with impressive names like artificial intelligence and machine learning. Certainly, there is much to be gained from prudent use of new predictive capabilities. However, the technologies come with significant risks of psychological harm, privacy violations, and discrimination, among others. Moreover, predictive models and learning algorithms are often flawed and produce erroneous outcomes. Many of these potential harms were previously considered and addressed in the context of genetic testing. Scientists and policy-makers would be wise to adopt similar approaches for long-term predictive health analytics. This paper has proposed just a few legal and non-legal interventions. These are designed to enhance data subjects' privacy rights, anti-discrimination protections, and ability to make informed decisions about obtaining disease predictions. However, many more minds must tackle the challenges of predictive health analytics and develop mechanisms to enhance the integrity and benefits of these technologies. Ignoring the potential perils and unintended consequences of long-term predictive analytics is imprudent and could cost society dearly.