



CSU  
College of Law Library

4-28-2021

# Unexpected Inequality: Disparate-Impact From Artificial Intelligence in Healthcare Decisions

Sahar Takshi

Follow this and additional works at: <https://engagedscholarship.csuohio.edu/jlh>



Part of the [Civil Rights and Discrimination Commons](#), [Health Law and Policy Commons](#), and the [Science and Technology Law Commons](#)

[How does access to this work benefit you? Let us know!](#)

### Recommended Citation

Sahar Takshi, *Unexpected Inequality: Disparate-Impact From Artificial Intelligence in Healthcare Decisions*, 34 J.L. & Health 215 (2021)  
available at <https://engagedscholarship.csuohio.edu/jlh/vol34/iss2/6>

This Article is brought to you for free and open access by the Journals at EngagedScholarship@CSU. It has been accepted for inclusion in Journal of Law and Health by an authorized editor of EngagedScholarship@CSU. For more information, please contact [library.es@csuohio.edu](mailto:library.es@csuohio.edu).

# UNEXPECTED INEQUALITY: DISPARATE-IMPACT FROM ARTIFICIAL INTELLIGENCE IN HEALTHCARE DECISIONS

Sahar Takshi, J.D.

*Systemic discrimination in healthcare plagues marginalized groups. Physicians incorrectly view people of color as having high pain tolerance, leading to undertreatment. Women with disabilities are often undiagnosed because their symptoms are dismissed. Low-income patients have less access to appropriate treatment. These patterns, and others, reflect long-standing disparities that have become engrained in U.S. health systems.*

*As the healthcare industry adopts artificial intelligence and algorithm-informed (AI) tools, it is vital that regulators address healthcare discrimination. AI tools are increasingly used to make both clinical and administrative decisions by hospitals, physicians, and insurers—yet there is no framework that specifically places nondiscrimination obligations on AI users. The Food & Drug Administration has limited authority to regulate AI and has not sought to incorporate anti-discrimination principles in its guidance. Section 1557 of the Affordable Care Act has not been used to enforce nondiscrimination in healthcare AI and is under-utilized by the Office of Civil Rights. State level protections by medical licensing boards or malpractice liability are similarly untested and have not yet extended nondiscrimination obligations to AI.*

*This Article discusses the role of each legal obligation on healthcare AI and the ways in which each system can improve to address discrimination. It highlights the ways in which industries can self-regulate to set nondiscrimination standards and concludes by recommending standards and creating a super-regulator to address disparate impact by AI. As the world moves towards automation, it is imperative that ongoing concerns about systemic discrimination are removed to prevent further marginalization in healthcare.*

## TABLE OF CONTENTS

I.	INTRODUCTION	217
II.	ARTIFICIAL INTELLIGENCE IN HEALTHCARE	219
	A. <i>How is Artificial Intelligence Used in the Healthcare Industry?</i>	219
	B. <i>UnitedHealth as a Cautionary Tale: Biases in AI and Resulting Discrimination</i>	221
III.	AGENCY OVERSIGHT OF HEALTHCARE AI	224
	A. <i>Food &amp; Drug Administration's Current Approach to AI</i>	224
	B. <i>Barriers and Improvements to FDA's Approach</i>	228
IV.	UNTESTED WATERS: THE ROLE OF NONDISCRIMINATION IN HEALTHCARE AI	231
	A. <i>Section 1557: Enforcement</i>	233
	B. <i>Section 1557: Private Right of Action</i>	239
V.	COMPLIANCE CHALLENGES	241
	A. <i>Licensing</i>	241
	B. <i>Malpractice Liability</i>	242
VI.	RECOMMENDATIONS	245
	A. <i>Industry Standards and Internal Compliance</i>	245
	B. <i>Recommended Regulations and a "Super Regulation"</i>	248
VII.	CONCLUSION	250

## I. INTRODUCTION

A 2019 study revealed that an algorithm used by UnitedHealth, one of the nation's largest managed care organizations, might be violating state and federal law: The algorithm had a racially discriminatory impact.<sup>1</sup> The algorithm (called "Impact Pro") makes eligibility determinations for "high risk care management" services by identifying patients with complex health needs.<sup>2</sup> The researchers found that it deemed black patients' health needs as "less than" white patients', and as a result, black patients were not targeted to benefit from specialized care management programs.<sup>3</sup> Such discriminatory effects from artificial intelligence and augmented intelligence (AI) are not undocumented;<sup>4</sup> however, the study was the first to expose these effects from automation in the healthcare industry. One can imagine an AI system that relies on a patient's oral description of their symptoms to design a treatment plan, or automated imaging technology that diagnoses skin conditions—both systems have the potential to discriminate against patients because it has been proven that AI systems have greater difficulty understanding African American vernacular and analyzing images of people of color.<sup>5</sup>

Discrimination in the healthcare industry is not a novel concept. Thirty-five years ago, then-Secretary Margaret Heckler issued a report and recommendations based on the findings from the Task Force on Black and Minority health, with the report's focal point being minority groups experiencing tremendous amounts of "excess deaths" compared to their non-minority counterparts.<sup>6</sup> Despite the Heckler

---

<sup>1</sup> Ziad Obermeyer et al., *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 336 SCIENCE 447, 447 (2019), <https://www.ehdc.org/sites/default/files/resources/files/Dissecting%20racial%20bias%20in%20an%20algorithm%20used%20to%20manage%20the%20health%20of%20populations.pdf>.

<sup>2</sup> *Id.*

<sup>3</sup> *Id.*; Isabel F. Farhi, *NY Action Against UnitedHealth Algorithm*, CONSUMER L. ROUND-UP (Nov. 12, 2019), <https://consumer.jenner.com/2019/11/ny-action-against-unitedhealth-algorithm.html>.

<sup>4</sup> Tom C.V. Lin, *Artificial Intelligence, Finance, and the Law*, 88 FORDHAM L. REV. 531 (2019) (financial industry); Charles A. Sullivan, *Employing AI*, 63 VILL. L. REV. 395 (2018) (employment); Karen Hao, *AI is sending people to jail—and getting it wrong*, MIT TECH. REV. (Jan. 21, 2019), <https://www.technologyreview.com/s/612775/algorithms-criminal-justice-ai/> (criminal justice system).

<sup>5</sup> See generally Cade Metz, *There Is a Racial Divide in Speech-Recognition Systems, Researchers Say*, N.Y. TIMES (Mar. 23, 2020), <https://www.nytimes.com/2020/03/23/technology/speech-recognition-bias-apple-amazon-google.html> (reporting that even the best performing AI system misunderstood 27% of words spoken by Blacks); Larry Hardesty, *Study finds gender and skin-type bias in commercial artificial-intelligence systems*, MIT NEWS (Feb. 11, 2018), [http://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212?mod=article\\_inline](http://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212?mod=article_inline).

<sup>6</sup> U.S. DEP'T OF HEALTH & HUMAN SERVS., REPORT OF THE SECRETARY'S TASK FORCE ON BLACK & MINORITY HEALTH (1985), <https://archive.org/stream/reportofsecretar00usde#page/10/mode/2up>.

Report's call to action—increased education and information, professional development, and research and data gathering—health disparities have persisted. Racial and ethnic minorities continue to experience higher rates of premature death and chronic disease. Native Americans and Alaskan natives have higher rates of infant mortality, and black patients are more likely to be inaccurately deemed as having high pain tolerance.<sup>7</sup> From a healthcare entity's perspective, healthcare AI presents a significant compliance challenge issue because of the risk of discrimination and relevant regulations (or lack thereof).

The introduction of AI-informed decision making into the healthcare sphere will continue to exacerbate many of these inequities, and possibly introduce new ones (e.g., in diagnosis and treatment decisions). The promise of AI as a more consistent, and even more accurate, decisionmaker means automation is likely to become the standard in healthcare, but should these positives outweigh its discriminatory impact? This Article proceeds in four parts. In Part I, this Article will outline the current and prospective uses of AI in healthcare and provide examples of potential discriminatory effects. Part II discusses the Food & Drug Administration's current efforts to regulate AI used in medical settings, particularly as clinical decision supports. This Part also highlights the gaps in regulations and makes recommendations to bolster the agency's role in fighting healthcare discrimination. Part III will introduce Section 1557 of the Affordable Care Act and argue that this nondiscrimination provision alone is inadequate to prevent or remedy disparate-impact from AI-informed decisions by providers and insurers beginning by describing the Department of Health and Human Service's enforcement of Section 1557, and further making recommendations for covered entities as they develop their compliance programs to address AI based on these enforcement actions. It then discusses the limited possibility of private rights of action for plaintiffs who are disparately impacted by healthcare AI. Part IV describes the novel compliance challenges posed by licensing laws and malpractice liability doctrines in relation to healthcare AI. Finally, Part V introduces recommendations for the healthcare industry to develop internal compliance standards and for regulators to promulgate policies that address biases in healthcare AI.

---

<sup>7</sup> NAT'L ACADS. OF SCIS., ENG'G, & MED., COMMUNITIES IN ACTION: PATHWAYS TO HEALTH EQUITY (James N. Weinstein et al. eds., 2017); Kelly M. Hoffman et al., *Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites*, 113 PROC. OF THE NAT'L ACAD. OF SCI. OF THE U.S. 4296, 4300 (2016), <https://www.pnas.org/content/pnas/113/16/4296.full.pdf>. The 2020 COVID-19 outbreak has exposed yet another pitfall in health equity in the United States, with many low-income individuals being unable to access treatment due to lack of insurance coverage and inability to practice social-distancing as a precautionary measure due to hourly employment without paid sick leave. See also Royce Dunmore, *How Coronavirus Affects Black People: Civil Rights Groups Call Out Racial Health Disparities*, NEWSONE (Mar. 13, 2020), <https://newsone.com/3911611/coronavirus-affects-black-people-civil-rights-groups-racial-health-disparities/>.

## II. ARTIFICIAL INTELLIGENCE IN HEALTHCARE

### A. *How is Artificial Intelligence Used in the Healthcare Industry?*

AI refers to a broad subset of computer sciences where machines are capable of making decisions that are typically made by humans.<sup>8</sup> Other industries use AI for decisions such as public-benefits eligibility determinations, risk–threat analysis, and employment recruitment efforts. In the healthcare industry, AI is increasingly being used for both administrative decisions and clinical decisions. Some examples include:

- Administrative decisions (e.g., making appointments, billing, reimbursement requests);
- Custodial (e.g., driverless vehicles to pull laundry, food services, clean rooms, automated pharmacy, cross-check travel conditions);
- Medical applications and wearables;
- Caregiving (“robotic” cribs, voice companions, electric lifts);
- Research and education;
- Clinical data analytics;
- Imaging, pathology, and radiology (e.g., detecting cancers, stroke, pneumonia, analyzing x-rays and scans);
- Predictive diagnosis (i.e., clinical decision supports); and
- Procedural AI (e.g., “tiny robots injected into the body for targeted drug delivery as an alternative to surgery”).<sup>9</sup>

AI is increasingly being used to make clinical determinations, such as to diagnose skin cancer<sup>10</sup> or recommend a combination of chemotherapy for cancer patients.<sup>11</sup> It can be used to determine individual patients’ risk of deteriorating, which allows physicians to predict which patients are likely to need to be transferred to the intensive care unit and intervene before a clinical emergency, increasing the rate of survival.<sup>12</sup> Tools like reSET-O (created by Pear Therapeutics)

---

<sup>8</sup> Press Release, Am. Med. Ass’n, AMA passes first policy recommendation on augmented intelligence (June 14, 2018), <https://www.ama-assn.org/press-center/press-releases/ama-passes-first-policy-recommendations-augmented-intelligence> (The American Medical Association’s 2018 policy recommendation prefers the term “augmented intelligence.”).

<sup>9</sup> Nicolas Terry, *Of Regulating Healthcare AI and Robots*, 21 YALE J. L. & TECH. 133, 144–47 (2019).

<sup>10</sup> Adewole S. Adamson & Avery Smith, *Machine Learning and Health Care Disparities in Dermatology*, 154 JAMA DERMATOLOGY 1247 (2018).

<sup>11</sup> See Robert Pearl, *Artificial Intelligence In Healthcare: Separating Reality From Hype*, FORBES (Mar. 13, 2018, 8:30 AM), <https://www.forbes.com/sites/robertpearl/2018/03/13/artificial-intelligence-in-healthcare/#7970d46b1d75>.

<sup>12</sup> See *id.*

treats opioid-use disorders with cognitive behavioral therapy through a mobile application.<sup>13</sup> Digital therapeutics created by Akili Interactive Labs work to treat or improvement cognitive impairments, such as ADHD, major depressive disorder, autism spectrum disorders, and multiple sclerosis through interactive digital therapies—similar to videogames.<sup>14</sup> Before proceeding, it is important to clarify that currently even when AI is used in healthcare for clinical purposes, “the physician, not the AI, has primacy.”<sup>15</sup>

Similarly, AI can be used to make administrative decisions outside of the examination or operating room. Managed care organizations use AI to prioritize risks in patients, allocate resources effectively, and allow physicians to intervene in patient care before health (and costs) skyrocket in critical situations.<sup>16</sup> AI can be used to automate medical billing—a task that is tedious, time consuming, and prone to errors when done manually.<sup>17</sup> AI also has the potential to further help providers make better clinical decisions—for example by informing physicians whether a patient is adhering to his or her therapy, and their response to that therapy.<sup>18</sup>

The risks of disparate-impact on suspect classes arising from AI-informed decision making should be of concern to healthcare providers, hospitals and clinics, and insurers. As these entities update their compliance and ethics programs to include factors such as privacy and fraud and abuse violations related to healthcare AI, they should also include nondiscrimination principles. Just like AI, an effective compliance program is a dynamic and constantly evolving system. The remainder of this Article will include discussions on how healthcare entities can incorporate nondiscrimination standards into their written policies and procedures, compliance auditing and investigation, training and education, and remediation procedures.

---

<sup>13</sup> Press Release, Pear Therapeutics, Sandoz Inc. and Pear Therapeutics Obtain FDA Clearance for reSET-O™ to Treat Opioid Use Disorder (Dec. 10, 2018), <https://peartherapeutics.com/sandoz-inc-and-pear-therapeutics-obtain-fda-clearance-for-reset-o-to-treat-opioid-use-disorder/>.

<sup>14</sup> *Getting to the root of cognitive impairment*, AKILI INTERACTIVE LABS, <https://www.akiliinteractive.com/science-and-technology> (last visited Mar. 11, 2020).

<sup>15</sup> Terry, *supra* note 9, at 137.

<sup>16</sup> Paul Nicolaus, *How Will AI Factor into the Future of Managed Care Markets*, 16 FIRST REPORT MANAGED CARE (Feb. 2019), <https://www.managedhealthcareconnect.com/articles/how-will-ai-factor-future-managed-care-markets>; Rachael Zimlich, *Real-World Applications of Artificial Intelligence in Healthcare*, MANAGED CARE EXECUTIVE (Oct. 24, 2018), <https://www.managedhealthcareexecutive.com/article/real-world-applications-artificial-intelligence-healthcare> (“[C]ertain artificial intelligence tools that interact with patients may also allow us to measure adherence and pill-taking behavior in real-time so we can identify changes in behavior immediately instead of after-the-fact.”).

<sup>17</sup> John Bailey, *Artificial Intelligence & Healthcare: Rethinking Medical Billing*, CHETU, <https://www.chetu.com/blogs/healthcare/ai-and-healthcare-medical-billing.php> (last visited Mar. 16, 2020).

<sup>18</sup> Zimlich, *supra* note 16.

*B. UnitedHealth as a Cautionary Tale: Biases in AI and Resulting Discrimination*

A recent study found that an algorithm used by U.S. health systems (UnitedHealth Group) has been discriminating against black patients. This algorithm, created by Optum, was used to identify the most high-risk patients to inform allocation of funds in the healthcare system. The algorithm used health care costs to make its predictions; however, spending for black patients is lower than for white patients due to “unequal access to care.”<sup>19</sup> These historic racial disparities in access to care translated into a racial bias in the algorithm—as a result, only 17.7% of black patients were identified as high-risk, but the study estimates that the true number should have been 46.5%.<sup>20</sup> In a letter to UnitedHealth, New York officials stated: “By relying on historic spending to triage and diagnose current patients, your algorithm appears to inherently prioritize white patients who have had greater access to healthcare than black patients.”<sup>21</sup> The racial bias in Optum’s algorithm not only presents a discrimination problem (in the form of disparate-impact), but can also harm individual patients in that it hinders physicians from not intervening in advance of a medical crisis.

Discrimination as a result of biases in AI is well-documented in other fields,<sup>22</sup> but the UnitedHealth case study is the only publicly available evidence of such effects in the health care context.<sup>23</sup> Scholars suspect that the discriminatory

---

<sup>19</sup> Obermeyer et al., *supra* note 1, at 447–48 (noting that while healthcare costs and medical needs are generally correlated (“sicker patients need and receive more care, on average”), deviations arise from health disparities: lack of healthcare insurance, resulting from geography, transportation, and job demands, and “taste-based” discrimination (e.g., white physicians not recommending preventative care to black patients)).

<sup>20</sup> *Id.* at 449.

<sup>21</sup> Letter from Linda Lacewell, Superintendent, N.Y. State Dep’t of Fin. Servs., and Howard A. Zucker, Comm’r, N.Y. State Dep’t of Health, to David S. Wichmann, CEO, UnitedHealth Group Inc. (Oct. 25, 2019), <https://www.dfs.ny.gov/system/files/documents/2019/10/20191025160637.pdf> (responding to the Obermeyer study and stating that UnitedHealth may not “produce, rely on, or promote an algorithm that has a discriminatory effect.”) [hereinafter *NY Letter to UnitedHealth*].

<sup>22</sup> See Soniya Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 U.C.L.A. L. REV. 54, 71 (2019) (LGBTQ+ individuals in targeted ads); McKenzie Raub, *Bots, Bias and Big Data: Artificial Intelligence, Algorithmic Bias and Disparate Impact Liability in Hiring Practices*, 71 ARK. L. REV. 529, 540–43 (2018) (women and racial minorities in hiring algorithms); Kristilian Lum & Wilian Isaac, *To predict and to serve?*, SIGNIFICANCE MAGAZINE 1 (Oct. 2016), [https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1740-9713.2016.00960.x?shared\\_access\\_token=IDAY641RsV7xsRV6B1hl04ta6bR2k8jH0KrdpFOx6C677Uo2ZpRJGfPu343uxAHkqNhmTYCA7Luw-6h3KOpptkCw153jVJOMXzw25e\\_4e82JLu75RVoBQza76WyvWfK7BJN2b0K0j-zwfXRgYNqUDbybnAVsHSC6KhR9WXrKYTMY%3D](https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1740-9713.2016.00960.x?shared_access_token=IDAY641RsV7xsRV6B1hl04ta6bR2k8jH0KrdpFOx6C677Uo2ZpRJGfPu343uxAHkqNhmTYCA7Luw-6h3KOpptkCw153jVJOMXzw25e_4e82JLu75RVoBQza76WyvWfK7BJN2b0K0j-zwfXRgYNqUDbybnAVsHSC6KhR9WXrKYTMY%3D) (racial bias in predictive policing).

<sup>23</sup> States, however, are aware of this effect. See, e.g., NY Letter to UnitedHealth, *supra* note 21, at 2 (noting that New York Governor Cuomo had already created a plan to combat racially



effects of AI will seep into the healthcare realm as well.<sup>24</sup> A major barrier to analyzing the frequency and magnitude of discrimination and physical harm resulting from AI is the proprietary nature of the algorithms.<sup>25</sup> Without more research, it is impossible to estimate the extent of biases in healthcare AIs outside of UnitedHealth; however, research from other industries is a good place to start.

The human-like decision making capacity of AI is developed by first introducing a training data-set—i.e., historical data that the AI uses to detect patterns and make future predictions.<sup>26</sup> One way that bias can enter into an AI system is unintentional lack of diversity in the training dataset. For example, visual AI tools have been shown to display racial and gender disparities: A facial recognition AI software was unable to accurately identify over one-third of black women in a photo lineup (notably, the algorithm was trained on a majority male and white dataset).<sup>27</sup> Voice-recognition systems show similar “race gaps,” tools created by companies such as Microsoft and Apple are unable to understand speech by black people.<sup>28</sup> The UnitedHealth case-study reflects how such invisible biases in training data can disparately impact certain groups; not only might some patients be excluded from special programs or preventative-measures, but physicians might fail to diagnose certain patients who have been under-diagnosed historically. The same problem is likely to arise when AI is trained not just on historical data, but on data from research. The fact that women and minorities are frequently excluded from medical research means that any AI based on such research may not be fully representative of healthcare consumers.<sup>29</sup>

---

disparities in mortality rates by creating the Maternal Review Board and making recommendations to the Department of Health).

<sup>24</sup> See Sharona Hoffman, *What Genetic Testing Teaches About Predictive Health Analytics*, 98 N.C. L. REV. 123 (2019) (suggesting that AI in genetic testing may erroneously conclude that some women are at a higher risk for various health problems and disproportionately identify patients with criminal records as being high risk).

<sup>25</sup> See Obermeyer et al., *supra* note 1, at 447.

<sup>26</sup> The potential biases discussed in this section are not only errors in the original dataset, they continue to amplify at every stage of the AI’s decision-making process. See Katyal, *supra* note 22, at 67–68 (“Errors at any stage can become amplified in the next stage, producing deviant outcomes in complex, troubling and sometimes difficult-to-detect ways.”).

<sup>27</sup> Hardesty, *supra* note 5. In fact, many of the AI systems that have been cleared by the Food & Drug Administration have been visual or scanning AI tools—e.g., Contact (identifies symptoms of stroke), IDx-DR (identifies diabetic retinopathy), and Accipio (identifies intracranial hemorrhage). See David Muoio, *Roundup: 12 healthcare algorithms cleared by the FDA*, MOBILE HEALTH NEWS (Nov. 15, 2018), <https://www.mobihealthnews.com/content/roundup-12-healthcare-algorithms-cleared-fda>.

<sup>28</sup> Metz, *supra* note 5 (“This indicates that the problem lies in the way the systems are trained to recognize sound.”).

<sup>29</sup> See Dhruv Khullar, *A.I. Could Worsen Health Disparities*, N.Y. TIMES (Jan. 3, 2019), <https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html>.

In the criminal-justice system, AI is used to determine the likelihood of crime in certain neighborhoods and the predictions are then used by police stations to intervene in crimes. Unsurprisingly, this use of AI can lead to confirmation bias that perpetuates racial disparities—police continue to target communities of color for drug use because the algorithm seems to be correct (despite the fact that drug use is equal across races).<sup>30</sup> The algorithm does not determine where crime is *actually* likely to occur, but rather where police is likely to *detect* the crime based on where they have detected them previously. These engrained stereotypes and the related confirmation bias can lead to similar challenges in the healthcare system—for example, an AI system might assume that low-income patients are less likely to adhere to their treatment plans, leading physicians not to focus on those patients.<sup>31</sup> Similarly, it could lead to overdiagnosis of conditions among certain populations,<sup>32</sup> which may also result in inequitable insurance rates.<sup>33</sup>

Depending on the use, the developers may *intentionally* exclude outlier data when they create AI. For example, Google may exclude or minimize a “minority interpretation of a search term” in its search results algorithm because it would not “help Google show relevant ads, generate clicks or produce revenue on a mass scale.”<sup>34</sup> One could easily imagine a healthcare system excluding an outlier patient, for example with a rare disability or underrepresented ethnic group, from its AI tool. What effect would this have when a future patient with those characteristics enters the hospital system? If the AI is a clinical tool, would it lead to a misdiagnosis because the AI could not detect the rare disease? If the AI is an administrative tool, would it lead to fewer resources being allocated to that patient because the AI underestimated the costliness of treating that patient?<sup>35</sup>

---

<sup>30</sup> See Lum & Isaac, *supra* note 22, at 5.

<sup>31</sup> Alternatively, given the recent shift towards value-based care, physicians might instead focus *more* on those patients. This dichotomy reflects a key challenge with AI in healthcare: Does the responsibility for and liability from decision-making lie with the AI or with the provider? This issue will be explored more in Part IV.B (Malpractice Liability).

<sup>32</sup> Effy Vayena, *Machine learning in medicine: Addressing ethical challenge*, 15 PLOS MED. (2018), <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002689> (“[I]n population data biased by the entrenched overdiagnosis of schizophrenia in African Americans.”).

<sup>33</sup> See Katyal, *supra* note 22, at 96 (noting that health insurers can charge higher premiums based on irrelevant characteristics).

<sup>34</sup> *Id.* at 69–70.

<sup>35</sup> Another form of discrimination can arise not from the AI making decisions, but the information that the AI collects which is later shared with other entities. Healthcare data in particular is very valuable. See CHRIS WAYMAN & NATASHA HUNERACH, *REALISING THE VALUE OF HEALTH CARE DATA: A FRAMEWORK FOR THE FUTURE 2* (2019), [https://assets.ey.com/content/dam/ey-sites/ey-com/en\\_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf](https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf) (noting that the 55 million patient records held by the UK’s National Health Service are valued at £5 billion, or just over \$6 billion USD). Providers and insurers have an incentive to sell the non-individually identifiable information that the AI learns through its decision-making for profit. Any

The possible biases described in this section reflect more than just errors in an algorithm, they indicate the potential for real harm to patients—such as physical harm (in the form of failed, delayed, or misdiagnosis, and inappropriate treatment) and social harm (in the form of inequitable billing, insurance rates, and coverage, and perpetuating stereotypes based on over-diagnosis). An analogy from Quartz magazine is apt: “Algorithms are like drugs,” they affect lives (i.e., have significant impact on the people/patients), perform differently on different demographics (i.e., algorithmic bias), and can have side effects (i.e., intended or unintended effects as a result of bias).<sup>36</sup> Unlike prescription drugs, however, healthcare AI is largely unregulated, meaning healthcare entities have few legal obligations to comply with (especially in terms of disparate-impact analysis) and patients have little legal recourse.

### III. AGENCY OVERSIGHT OF HEALTHCARE AI

#### A. Food & Drug Administration’s Current Approach to AI

Despite the increased use of healthcare AI, there is little regulation of the emerging technology. Federal agencies, state medical boards, and trade organizations have only just begun to address the fate of this new technology. The Food & Drug Administration (FDA) has authority to regulate a medical device, which is “an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar related article” recognized as a pharmaceutical, used for diagnosis or treatment, or intended to affect the structure or function of the body.<sup>37</sup> The 21<sup>st</sup> Century Cures Act, significantly limited the FDA’s ability to regulate AI; devices that support or provide recommendations to a healthcare professional, and the healthcare professional has the opportunity to “independently review the basis for such recommendation” are excluded from the definition of device.<sup>38</sup> Notably, the FDA does not have the statutory authority to regulate devices

---

biases that exist in that data can result in discrimination in whatever context the sold data is used (regardless of whether it resulted in discrimination in the AI’s original purpose). For example, an algorithm that over-estimated patients’ propensity for a certain illness at a community hospital, if sold, may result in life-insurance carriers charging higher premiums for people from that community. Although this paper will not discuss this in depth, the GDPR model used in European nations could protect against this form of second-degree discrimination. *See* Katyal, *supra* note 22, at 106–07 (noting that the GDPR model “requires individuals to have the right to confirm whether their personal data is being processed, the purpose of the process, the source of the data, and the logic behind any automated process”).

<sup>36</sup> Andy Coravos et al., *We should treat algorithms like prescription drugs*, QUARTZ (Feb. 14, 2019), <https://qz.com/1540594/treating-algorithms-like-prescription-drugs-could-reduce-ai-bias/>.

<sup>37</sup> 21 U.S.C. § 321 (2018).

<sup>38</sup> 21<sup>st</sup> Century Cures Act, Pub. L. No. 114-225, § 3060(a), 130 Stat. 1033 (codified as amended at 21 U.S.C. § 520(o)(E)(ii)–(iii)). Specifically, the Act excludes devices that meet these criteria from the definition of medical device: (1) not intended to analyze medical images or signals from in vitro diagnostic devices or analyze a pattern from signal acquisition systems; (2) intended to be

that are used as “administrative support[s],” such as billing and analyzing population level data.<sup>39</sup> The FDA follows different processes to provide clearance or approval prior to marketing of a device based on its class;<sup>40</sup> medical devices are reviewed under different standards (e.g., premarket clearance (510(k)), De Novo classification, or premarket approval) that highlight the risk to patients.<sup>41</sup>

The agency’s regulatory framework, created in a pre-AI era, does not directly address AI technology, but the agency has developed guidance and a proposed a framework in the last two years. The FDA has previously regulated software as a medical device (SaMD)—i.e., “software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device.”<sup>42</sup> Examples of SaMD include fertility apps that track reproductive outcomes and software that allows physicians and patients to view MRI results on smartphones.<sup>43</sup> The FDA has issued guidance regulating SaMD, and a proposed regulatory framework for AI and machine-learning SaMDs.<sup>44</sup> These

---

used to display, analyze or print patient medical information (or information from clinical studies); (3) intended to support or give recommendations to healthcare professionals about preventing, diagnosing, or treating diseases; and (4) intended to be independently reviewed by a healthcare professional (as opposed to professional primarily relying on the recommendations).

<sup>39</sup> See Food, Drug & Cosmetic Act (FDCA) of 1938, 21 U.S.C. § 360j(o)(1)(A) (providing in relevant part, “[t]he term device . . . shall not include a software function that is intended for administrative support of a health care facility, including the processing and maintenance of financial records, claims or billing information, appointment schedules, business analytics; information about patient populations, admissions, practice and inventory management, analysis of historical claims data to predict future utilization or cost-effectiveness, determination of health benefit eligibility, population health management, and laboratory workflow . . .”).

<sup>40</sup> Floor Van Leeuwen, *A 101 guide to the FDA-regulatory process for AI healthcare software* (Nov. 20, 2019), <https://www.quantib.com/blog/a-101-guide-to-the-fda-regulatory-process-for-ai-radiology-software> (explaining that class I and II are relatively low-patient risk products, whereas class III are products that sustain or support life, are implanted, or present a potential unreasonable risk of injury to the patient) [hereinafter *FDA Regulatory Process for AI*].

<sup>41</sup> *Artificial Intelligence and Machine Learning in Software as a Medical Device*, FDA.GOV, <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (last visited Jan. 28, 2020).

<sup>42</sup> *Id.* (explaining that the Federal Food, Drug, and Cosmetic Act defines medical purpose as those intended to treat, diagnose, mitigate, or prevent disease or other conditions).

<sup>43</sup> See Evan Heier, *SaMD: Everything You Need to Know About Software as a Medical Device*, SELECTHUB (Sept. 30, 2019), <https://www.selecthub.com/medical-software/software-medical-device-samd/> (noting that software that “principally drive[s] a hardware device,” such as the software that enables the MRI to work, is not SaMD).

<sup>44</sup> U.S. FOOD & DRUG ADMIN., SOFTWARE AS A MEDICAL DEVICE (SAMD): CLINICAL EVALUATION GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF (2017), <https://www.fda.gov/media/100714/download> [hereinafter *SAMD GUIDANCE*]; U.S. FOOD & DRUG ADMIN., PROPOSED REGULATORY FRAMEWORK FOR MODIFICATIONS TO ARTIFICIAL INTELLIGENCE/MACHINE LEARNING [AI/ML]-BASED SOFTWARE AS A MEDICAL DEVICE [SAMD] DISCUSSION PAPER AND REQUEST FOR FEEDBACK, <https://www.fda.gov/media/122535/download>

documents focus on SaMD manufacturer’s responsibility to test for the validity of clinical associations (i.e., accuracy, reliability, and precision) and provide a detailed framework for evaluating for safety, effectiveness and performance; yet these documents do not provide robust guidance (or create obligations) for detecting and remedying discrimination.<sup>45</sup>

In 2019, the FDA’s Digital Health program released the Digital Health Innovation Action Plan.<sup>46</sup> The Plan recognized that

FDA’s traditional approach . . . is not well suited for the faster iterative design, development, and type of validation used for software-based medical technologies. Traditional implementation of the premarket requirements may impede or delay patient access to critical evolutions of software technology, particularly those presenting a lower risk to patients.<sup>47</sup>

Notably, the Digital Health Program developed policies and approaches to balance the benefits and risks to patients—one such approach was to not enforce compliance for “low risk” technologies so that they are readily available to consumers.<sup>48</sup> This begs the question: Did the FDA consider disparate-impact on vulnerable populations in making such risk assessments? Most importantly, the FDA stated its intent to develop a pre-certification program that “could replace the need for a premarket submission for certain products and allow for decreased submission content and/or faster review of the marketing submission for other products.”<sup>49</sup> The pre-cert program would regulate the *developer* rather than the product. The agency could pre-certify developers that “demonstrate a culture of quality and organizational excellence based on objective criteria.”<sup>50</sup> Such criteria can include excellent software design, development, or validation and testing, and the pre-certified developers could market low-risk devices without additional review (or at least less rigorous review). FDA accepted nine companies to

---

(adopting international industry principles created by the International Medical Device Regulators Forum (IMDRF)).

<sup>45</sup> SAMD GUIDANCE, *supra* note 44, at 4, 7, 12. For example, the Guidance requires clinical evaluation to be “iterative and continuous” and include independent review.

<sup>46</sup> U.S. FOOD & DRUG ADMIN., DIGITAL HEALTH INNOVATION ACTION PLAN (2019), <https://www.fda.gov/media/106331/download>.

<sup>47</sup> *Id.* at 2.

<sup>48</sup> *Id.* at 2–3 (stating that FDA did not focus oversight or enforce compliance for low-risk mobile apps; technologies that only transmit, store, or data; and products that only promote general wellness).

<sup>49</sup> *Id.* at 5.

<sup>50</sup> *Id.*

participate in the Pre-Cert pilot program in August 2017.<sup>51</sup> It stated in the NPRM that one factor in eligibility for the program is measured on “Key Performance Indicators” or similar measures but made no specific mention of nondiscrimination measures.

Recently, the FDA released a draft guidance around Clinical Decision Support (CDS) software—software that includes computerized alerts for patients and providers, condition-specific order sets, diagnostic support, and more.<sup>52</sup> The CDS guidance specifically applies to such software that make patient-specific recommendations to a health care provider, such as possible diagnosis, recommended treatment plans, and recommended diagnostic tests.<sup>53</sup> Key to this definition is that the CDS function must *not* be independently reviewed by a healthcare provider (labelled “Device CDS”).<sup>54</sup> A healthcare AI can escape FDA enforcement under the guidance (labelled “Non-Device CDS”) if the manufacturer provides a plain-language description of the software (including the intended use and intended user), the data inputs required to generate a recommendation (e.g., patient’s age), and the basis for the recommendation.<sup>55</sup> To adequately describe the basis-for-recommendation, the guidance requires software developers to “describe the underlying data used to develop the algorithm and [] include plain language descriptions of the logic or rationale used by the algorithm to render a recommendation.”<sup>56</sup> Even for those Device CDSs that the FDA has authority to regulate, the agency stated it intends to focus enforcement efforts on “serious or critical situations or conditions” where the provider is unable to independently review the Device CDS’s recommendation—e.g., machine-learning algorithms that categorize symptoms of flus, software that identifies signs of opioid addiction, and

---

<sup>51</sup> Fostering Medical Innovation: A Plan for Digital Health Devices; Software Precertification Pilot Program, 82 Fed. Reg. 35,216, 35,217 (July 28, 2017), <https://www.govinfo.gov/content/pkg/FR-2017-07-28/pdf/2017-15891.pdf> (noting that nine companies were selected to participate in the pilot: Apple, Fitbit, Johnson & Johnson, Pear Therapeutics, Phosphorus, Roche, Samsung, Tidepool, and Verily); *Digital Health Software Precertification (Pre-Cert) Program*, FDA.GOV, <https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-precert-program> (last visited Mar. 10, 2020).

<sup>52</sup> U.S. FOOD & DRUG ADMIN., CLINICAL DECISION SUPPORT SOFTWARE DRAFT GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF 5 (Sept. 27, 2019), <https://www.fda.gov/media/109618/download> [hereinafter CDS DRAFT GUIDANCE]. The guidance clarifies that CDS are different from SaMD in that they are only intended to be used as a support by providing recommendations, rather than providers relying primarily on the CDS recommendations.

<sup>53</sup> *Id.* at 6–7, 11. In line with the 21st Century Cures Act, the CDS guidance addresses software that is not excluded from the definition of medical device (i.e., software that provides health recommendations about prevention, diagnoses, or treatment *without* an opportunity for independent review by a physician).

<sup>54</sup> See 21<sup>st</sup> Century Cures Act, *supra* note 38 and accompanying text.

<sup>55</sup> CDS DRAFT GUIDANCE, *supra* note 52, at 12.

<sup>56</sup> *Id.*

machine learning algorithms that identify diabetics at risk of post-operative cardiovascular difficulties.<sup>57</sup>

### *B. Barriers and Improvements to the FDA's Approach*

The nature of the FDA's oversight of AI technology creates certain compliance challenges for providers and AI developers, and leaves patients at risk of discrimination. First, the lack of enforced regulations or guidance at this time means there is little accountability for the AI developers to ensure that algorithms and software do not perpetuate biases.<sup>58</sup>

Second, the patchwork of guidance and proposed regulatory frameworks leaves several gaps in the enforcement around AI. Most importantly, FDA regulation authority is limited to medical devices—those that play a role in diagnosis and treatment of disease—they can never extend to AI used for administrative purposes, which can have an equally discriminatory impact on healthcare consumers.<sup>59</sup> This caveat in the statutory language leaves a large hole in terms of enforcement authority; absent amendments to the statute, the FDA will never be able to regulate biases in insurance-rating AI or AI used to identify social determinants of health.

Furthermore, the added caveats introduced by the 21<sup>st</sup> Century Cures Act leave a large portion of healthcare AI technology unregulated. Per the Act and the FDA's draft guidance document, Non-Device CDSs are those technologies where physicians maintain primacy and manufacturers ensure transparency (e.g., by describing the algorithm's rationale).<sup>60</sup> As discussed in Part IV, a method of countering AI bias is to require information about its design, testing, and rationale to be available to providers and patients.<sup>61</sup> There is clearly tension between the FDA's approach and what researchers and advocates agree is a necessary step to limit the discriminatory effects of AI bias. It seems that the FDA will not be able to regulate healthcare AI if the manufacturers provide even limited opportunity for

---

<sup>57</sup> *Id.* at 14, 23. Similarly, it will focus regulatory oversight over CDS's intended for patients that conducts continuous glucose monitoring, assess patient's stress/anxiety to provide treatment recommendations, or providers recommendations to caregivers of children with cystic fibrosis about when to bring the child to the ER.

<sup>58</sup> See generally *id.* at 14 (noting at the top of the CDS guidance document “Contains *Nonbinding* Recommendations”) (emphasis added).

<sup>59</sup> Food, Drug & Cosmetic Act (FDCA) of 1938, 21 U.S.C. § 360j(o)(1)(A).

<sup>60</sup> CDS DRAFT GUIDANCE, *supra* note 52, at 12.

<sup>61</sup> See IEEE, *infra* note 151 and accompanying text.

physicians to exercise independent judgement by giving cursory descriptions of the AI rationale.<sup>62</sup>

The FDA also specifically noted in its draft CDS guidance that it does not intend to enforce compliance for devices that inform providers for “non-serious situations or conditions” such as recommendations for allergens and common cold symptoms, alerts for cholesterol management, or recommendations of over-the-counter drugs.<sup>63</sup> Even for those Device CDSs that the FDA *can* regulate, it will only regulate a small percentage. Although the “serious situations and conditions” that the FDA plans to focus its attention present a higher risk of medical harm, even low-risk conditions have a high likelihood of discrimination if AI is used. Bias does not distinguish between high-risk and low-risk diseases, even something as simple as an AI that recommends over-the-counter medication can cause unnecessary pain if the AI tends to not recommend the medication to certain individuals.

Lastly, none of the FDA’s proposed frameworks adequately address the potential for disparate impacts on certain groups, which is a missed opportunity to ensure that AI *developers* incorporate nondiscrimination elements (e.g., through feedback loops) at the design and production stage. In the limited instances where the FDA does plan to regulate Device CDS, it is likely that its focus will be primarily on safety and effectiveness rather than AI bias.<sup>64</sup> While the Digital Health Innovation Action Plan seeks a “culture of quality” among AI developers in the proposed pre-certification program, it still fails to specifically address discrimination.<sup>65</sup> The FDA should include in its pre-certification criterion a requirement that manufacturers must attest to factors like commitment to nondiscrimination, review of training data for bias, and testing for disparate-impact.

Not only is there a regulatory gap, but also a compliance gap. Ordinarily, organizations develop their compliance programs around federal and state regulations; where technology is just emerging and the government has not yet caught up, organizations may find it difficult to identify the appropriate standards to which they should hold themselves. Similar dilemmas are arising in other industries where AI is emerging, such as the criminal justice system. California recently introduced a bill to regulate the use of AI in pretrial risk-assessment tools used by courts in place of bond hearings that received responses from several advocacy groups with recommendations of specific measures to reduce the risk of

---

<sup>62</sup> See CDS DRAFT GUIDANCE, *supra* note 52, at 12 (providing an example that a physician would be unable exercise independent judgement where the information provided by the manufacturer “could not be expected to be independently understood”).

<sup>63</sup> *Id.* at 20–21. Similarly, it does not intend to enforce regulations against devices where the patient has primacy to make decisions (such as to take allergy medication recommended by the CDS).

<sup>64</sup> See generally Elias Mallis, *An Introduction to FDA’s Regulation of Medical Devices*, FDA.GOV, <https://www.fda.gov/media/123602/download> (last visited Apr. 2, 2020).

<sup>65</sup> DIGITAL HEALTH INNOVATION ACTION PLAN, *supra* note 46, at 5.



bias.<sup>66</sup> In the absence of regulation around healthcare AI, providers, hospitals, and AI developers should take a cautious approach—limit AI to technology that has been tested for equality and its use to situations where risk of bias is lowest.

Healthcare AI instruments should be designed so that no patient of a suspect class is unduly burdened by its errors. There are several measures of fairness that the FDA should require AI developers to use during design and testing, including error rate balance and predictive parity.<sup>67</sup> The agency should look for a combination of such measures that is most apt to avoid discrimination.<sup>68</sup> The FDA, and AI developers and healthcare entities, should remain open to input from the community (such as advocacy groups) and independent data to identify the best fairness measures.<sup>69</sup> In the context of pretrial risk-assessment tools, advocacy groups in this context argue that “the design of any tool should always give far greater weight to the avoidance of false positives than to the avoidance of false negatives” because the detriment of erroneously detaining a person is worse than erroneously releasing one.<sup>70</sup> In the context of healthcare AI, however, the risks are less clear—is it a greater threat to the health and safety of patients to not diagnose an illness at all or to misdiagnose it as something else? Medical professionals should carefully consider and balance these risks to inform AI design and organization policies and procedures to ensure that risks (particularly risks arising out of bias) are mitigated to the fullest extent possible. Relatedly, the datasets used to train AI must be vetted for correctness and reliability.

Advocacy groups in the risk-threat context suggest “transparent and periodic examination” of relevant factors—in the healthcare context, this might include review of timely diagnoses, treatment success and adherence, and rehospitalization rates by race, age, gender, and disability.<sup>71</sup> Although there are currently no regulations, AI developers and healthcare entities can develop their own internal criterion to achieve transparency.<sup>72</sup> Implementing such internal

---

<sup>66</sup> See THE USE OF PRETRIAL “RISK ASSESSMENT” INSTRUMENTS: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS 3, <http://civilrightsdocs.info/pdf/criminal-justice/Pretrial-Risk-Assessment-Full.pdf> (last visited Mar. 23, 2020) [hereinafter STATEMENT ON RISK ASSESSMENT TOOLS].

<sup>67</sup> See *id.* at 3. Other methods of algorithmic fairness include: Fairness through blindness (i.e., exclusion of factors related to protected class); Group Fairness; Statistical Parity; Equal Group Error Rates (the rate of false negatives and positives is the same for all groups); Individual Fairness (same outcome regardless of patient’s group); Predictive Parity (equalized positive predictive values); and Similarity Measures (classifications based on similar characteristics relevant to a particular task). See Mark MacCarthy, *Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms*, 48 CUMB. L. REV. 67, 91 (2017–18).

<sup>68</sup> See STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 3.

<sup>69</sup> See *id.*

<sup>70</sup> See *id.* at 4.

<sup>71</sup> See *id.* at 4–5.

<sup>72</sup> See *id.* at 7. These factors will be discussed in Part IV Recommendations.

controls will ensure that the healthcare industry is prepared to immediately adapt to any regulations promulgated by the FDA (or other agencies) in the future.

#### IV. UNTESTED WATERS: THE ROLE OF NONDISCRIMINATION IN HEALTHCARE AI

Discriminatory effects from AI in other industries may be governed by federal nondiscrimination laws. For example, administrative agencies are increasingly using AI to make eligibility determinations for public benefits.<sup>73</sup> Where constitutionally protected property interests are at stake, the use of AI may present a procedural due process problem due to the lack of transparency (which affects the meaningful opportunity to challenge).<sup>74</sup> Where AI intentionally treats persons differently based on factors such as race or gender, the Equal Protection Clause may apply.<sup>75</sup> Similarly, facially neutral policies that disparately impact certain populations may run afoul the disparate-impact provisions of civil rights statutes. For example, employers frequently use machine-learning tools for resume screening, hiring and retention decisions, and identifying employees for promotions. Where employees and prospective employees have been discriminated against by these tools, they may pursue specific causes of action per disparate-impact (and business necessity) principles of Title VII.<sup>76</sup>

These remedies, however, do not extend to any discrimination resulting from the use of AI in the healthcare context. There is no constitutional property right to receive, for example, a correct diagnosis or specific treatment plan. Similarly, there is no such right to coverage for certain services and items or to be selected to participate in special health programs. Thus, when covered entities use AI to diagnose, treatment, and monitor patients, or in billing and utilization review,

---

<sup>73</sup> Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision-Making In The Machine-Learning Era*, 105 GEO. L.J. 1147 (2017).

<sup>74</sup> *But see id.* at 1186 (arguing that the use of algorithms in public benefits determinations is actually *more* likely to survive the *Mathews* balancing-test for procedural due process because it is perceived to reduce human prejudice).

<sup>75</sup> However, facially neutral laws or regulations that have a disparate impact on suspect classes are insufficient to trigger the Equal Protection Clause. *See* Interview by Leslie Garfield Tenzer with Emily Gold Waldman, Dean, Pace L. Sch. (Apr. 16, 2019), <https://lawtofact.buzzsprout.com/138309/1027580-the-equal-protection-clause-and-> (explaining that there needs to be both disparate impact and a showing of intent behind a facially neutral law to trigger heightened review).

<sup>76</sup> Raub, *supra* note 22, at 544–53. Moreover, federal agencies using AI could also pose a nondelegation issue under Article I of the Constitution. *See generally* Coglianese & Lehr, *supra* note 73, at 1178 (noting that, although broad delegation to administrative agencies has been a long-accepted practice of Congress, “the law has always assumed that the recipient of that authority would be a human being”).

constitutional protections may not apply.<sup>77</sup> The lack of protections and remedies for affected patients is not only due to the fact that our “statutory and constitutional schemes are poorly crafted to address issues of private, algorithmic discrimination” but also because the black-box effect created by the proprietary nature of such algorithms makes it difficult for patients to discover such effects.<sup>78</sup>

While most aspects of healthcare are not constitutionally protected, one federal statute does shield patients from discrimination: Section 1557 of the Patient Protection and Affordable Care Act (ACA).<sup>79</sup> Section 1557 is the nondiscrimination provision that prohibits certain health systems from discriminating against patients on the basis of race, sex, national origin, disability and other characteristics.<sup>80</sup> It provides that:

[A]n individual shall not . . . be excluded from participation in, be denied the benefits of, or be subjected to discrimination under, any health program or activity, any part of which is receiving Federal financial assistance, . . . or under any program or activity that is administered by an Executive Agency or any entity established under this title.<sup>81</sup>

Section 1557 applies to *any* health program or activity, any part of which receives federal financial assistance, is administered by a federal agency, or is established under the ACA—this includes not only physicians and hospitals, but

---

<sup>77</sup> Katyal, *supra* note 22, at 99 (“Like the civil rights era that came before it, AI is implicated within a vast array of decisions that come, not from the government, but from the private sector, even if many of them implicate civil rights in the process. For example, the right to be considered for employment, free from consideration of one’s disability—the right at issue in the Kyle Behm case just discussed—directly correlates to the right to work. Similarly, the right to an education, the right to vote, the right to make contracts, the right to travel, the right to get insurance, and the right to receive information, among others, are all at issue when an algorithm makes its (private) decisions about who does and who does not receive the entitlement and the conditions attached to it. Those decisions are not always subject to public oversight. And even more problematically, they may be shielded from view, due to trade secrecy and systemic opacity.”).

<sup>78</sup> *See id.* at 100 (noting further that these effects from the lack of transparency pose issues in non-private industries as well, for example discovering Fourth Amendment violations when AI is used as a predictive policing tool).

<sup>79</sup> *See* 42 U.S.C. § 18116(a) (2018). While other federal nondiscrimination statutes exist, they are not sweeping enough to cover all discrimination in all industries. *See, e.g., Doe v. BlueCross BlueShield of Tennessee*, 926 F.3d 235 (6th Cir. 2019) (noting in dicta that the Americans with Disabilities Act may not actually protect those covered by private insurance). For a discussion on why Title VI of the Civil Rights Act has been ineffective in preventing discrimination by federally funded healthcare entities, *see* Sidney D. Watson, *Section 1557 of the Affordable Care Act: Civil Rights, Health Reform, Race and Equity*, 55 HOWARD L.J. 855, 860–70 (2012).

<sup>80</sup> 42 U.S.C. § 18116(a).

<sup>81</sup> *Id.*

also insurers.<sup>82</sup> These entities are prohibited from excluding an individual from participating in or receiving the benefits, or otherwise discriminating against an individual, on the grounds prohibited under: (1) Title VI of the Civil Rights of 1964, (2) Title IX of the Education Amendments of 1972, (3) the Age Discrimination Act of 1975, or (4) the Rehabilitation Act of 1973. Stated simply, the statute and regulations implementing Section 1557 prohibit health entities from discriminating against patients on the basis of race, ethnicity, national origin, sex, age, or disability.

#### *A. Section 1557: Enforcement*

The Office of Civil Rights (OCR) within the U.S. Department of Health and Human Services is tasked with implementing and enforcing Section 1557. OCR receives complaints about Section 1557 violations and conducts “compliance reviews” to investigate such discrimination claims.<sup>83</sup> If a compliance review reveals noncompliance (that is within OCR’s jurisdiction), OCR first attempts to reach a voluntary agreement to remedy the discrimination.<sup>84</sup> However, if the covered entity does not voluntarily resolve the issue in a satisfactory manner, OCR will issue a Letter of Findings to outline why the entity is noncompliant and OCR’s next steps (including referral to DOJ for enforcement steps to terminate federal financial assistance).<sup>85</sup> Since the implementation of Section 1557, there have only been fifty-eight publicly available resolutions.<sup>86</sup> A few examples:

- Touro Infirmary Emergency Department of Louisiana (sex discrimination): OCR investigated a claim after a patient alleged he was denied appropriate care and treatment after a domestic violence incident, and was subject to rude comments based on his gender (male).<sup>87</sup>

---

<sup>82</sup> See OFFICE FOR CIVIL RIGHTS, *Section 1557: Coverage of Health Insurance in Marketplaces and Other Health Plans*, HHS.GOV, <https://www.hhs.gov/civil-rights/for-individuals/section-1557/fs-health-insurance/index.html> (last visited Aug. 25, 2016). Not covered, however, are third party administrators of self-insured plans. See 81 Fed. Reg. 31,376, 3,1432 (May 18, 2016) (“Third party administrators are generally not responsible for the benefit design of the self-insured plans they administer and that ERISA (and likely the contracts into which third party administrators enter with the plan sponsors) requires plans to be administered consistent with their terms.”).

<sup>83</sup> See Exec. Order No. 13,160, 65 Fed. Reg. 39,775 (June 23, 2000), *reprinted in* 42 U.S.C. 2000d (complaint procedures); see also OFFICE FOR CIVIL RIGHTS, *How OCR Enforces Civil Rights Discrimination Laws and Regulations*, HHS.GOV, <https://www.hhs.gov/civil-rights/for-providers/compliance-enforcement/enforcement-process/index.html> (last visited Sept. 29, 2015).

<sup>84</sup> *Id.*

<sup>85</sup> *Id.*

<sup>86</sup> For a list of recent Resolution Agreements by OCR, see OFFICE FOR CIVIL RIGHTS, *Recent Civil Rights Resolution Agreements & Compliance Reviews*, HHS.GOV, <https://www.hhs.gov/civil-rights/for-providers/compliance-enforcement/agreements/index.html> (last visited Mar. 2, 2020).

<sup>87</sup> See OFFICE FOR CIVIL RIGHTS, *OCR Enforcement under Section 1557 of the Affordable Care Act Sex Discrimination Case*, HHS.GOV, [https://www.hhs.gov/civil-rights/for-individuals/section-](https://www.hhs.gov/civil-rights/for-individuals/section-1557/fs-health-insurance/index.html)

- Pennsylvania Department of Human Services (PDHS) (discrimination on the basis of nationality—limited English proficiency): OCR investigated two complaints regarding access to programs in the PDHS Office of Income Maintenance by individuals with limited English proficiency (LEP) because they did not receive appropriate language services and were thus hindered from obtaining benefits in a timely manner.<sup>88</sup>
- Mid-Maryland Muskoskeletal Institute (MMI) (disability discrimination): OCR received a complaint that MMI violated Section 1557 by discriminating against the complainant and her minor son on the basis of their deafness because MMI failed to provide a qualified sign language interpreter.<sup>89</sup>
- Office of African American Children's Services (OAACS) (race discrimination): OCR received a complaint regarding OAACS's use of "racial classifications as the sole factor in determining which children" in a region of Washington state received child protective and welfare services in violation of Title VI of the Civil Rights Act.<sup>90</sup>

Like most healthcare statutes, Section 1557 presents a significant compliance responsibility for covered entities. OCR has not yet investigated any discrimination complaints regarding the use of AI by a covered entity, nor has it issued any guidance about how it might treat discrimination that originates from AI. Absent such information, covered entities can use the existing Voluntary Resolution Agreements as a starting point when developing their compliance programs.

As AI continues to permeate the healthcare industry, covered entities will need to revise their written policies and procedures to ensure that their existing nondiscrimination policies extend to AI usage. For example, if AI systems are used to identify certain patients as being victims of domestic violence, there should be

---

1557/ocr-enforcement-section-1557-aca-sex-discrimination/index.html (last visited Aug. 1, 2016) (sex discrimination) [hereinafter OCR Enforcement of Touro Infirmary].

<sup>88</sup> VOLUNTARY RESOLUTION AGREEMENT BETWEEN THE U.S. DEP'T OF HEALTH & HUMAN SERVS. OFFICE FOR CIVIL RIGHTS AND PENNSYLVANIA DEP'T OF HUMAN SERVS. 2 (2019), <https://www.hhs.gov/sites/default/files/hhs-padhs-vra.pdf> (discrimination based on race, ethnicity or national origin) [hereinafter OCR ENFORCEMENT OF PDHS].

<sup>89</sup> VOLUNTARY RESOLUTION AGREEMENT BETWEEN THE U.S. DEP'T OF HEALTH & HUMAN SERVS. OFFICE FOR CIVIL RIGHTS AND MID-MARYLAND MUSKOSKELETAL INSTITUTE 2 (2019), <https://www.hhs.gov/sites/default/files/MMI-vra.pdf> (discrimination on the basis of disability) [hereinafter OCR ENFORCEMENT OF MMI].

<sup>90</sup> VOLUNTARY RESOLUTION AGREEMENT BETWEEN THE U.S. DEP'T OF HEALTH & HUMAN SERVICES OFFICE FOR CIVIL RIGHTS, THE ADMIN. FOR CHILDREN & FAMILIES, AND THE WASHINGTON STATE DEP'T OF SOCIAL & HEALTH SERVS. 1 (2010), [https://www.hhs.gov/sites/default/files/oaacs\\_ra.pdf](https://www.hhs.gov/sites/default/files/oaacs_ra.pdf) (discrimination based on race, ethnicity, or national origin) [hereinafter OCR ENFORCEMENT OF OAACS]. OCR received this complaint prior to the passage of Section 1557, however, the agency uses this as an example of how it enforces the regulation for instances of race-based discrimination.

policies to ensure that the system’s identification is gender neutral.<sup>91</sup> Similarly, covered entities can develop criteria for their AI that reflect nondiscrimination principles and ensure that the software they purchase, create, or commission meets their requirements. For example, covered entities can create assessment criteria for AI that is used to identify patients who would benefit from auxiliary aids for their disabilities (e.g., to ensure those with rare disabilities are not excluded) or from interpretation services because they are limited English proficient (e.g., to ensure that those who speak less common languages can still access an interpreter).<sup>92</sup> Entities must also consider designating a nondiscrimination coordinator specifically for their AI usage to effectuate and oversee these policies.<sup>93</sup>

Importantly, covered entities should continue to monitor AI determinations—in both clinical and administrative purposes—to identify patterns of disparate impact. Review of AI decisions should be incorporated into entities’ scheduled audit plans and should be done regularly. In addition to looking factors such as data-privacy or error-rates, these audits must include a review of discrimination patterns.<sup>94</sup> To protect against confirmation-bias by covered entities’ officers and executives, they should also hire external auditors.<sup>95</sup> Third-party contractors will be particularly useful for auditing AI usage because the entities’ internal officers likely do not possess the requisite computer-science background to review the source codes to identify which data elements are contributing to the disparate impact. Monitoring and auditing alone is of course insufficient; covered entities must also develop strategic plans to timely remedy any biases they discover and should cease use of the AI until the discriminatory effects can be eliminated with some certainty.<sup>96</sup> Not only will this internal-remedy approach reflect covered

---

<sup>91</sup> See, e.g., OCR Enforcement of Touro Infirmary, *supra* note 87 (requiring Touro Infirmary Emergency Department of Louisiana to revise its protocol to reflect-gender neutral procedures around domestic violence incidents).

<sup>92</sup> See, e.g., OCR ENFORCEMENT OF MMI, *supra* note 89, at 9 (requiring the Institute to use specific factors created by OCR to prioritize patients requiring auxiliary aids and qualified interpreters).

<sup>93</sup> See, e.g., OCR ENFORCEMENT OF PDHS, *supra* note 88, at 7 (requiring the Pennsylvania agency to designate a “language assistance coordinator” in light of an investigation). See also 42 C.F.R. § 92.7(a) (2019) (stating that covered entities of more than fifteen employees are already required to designate someone to oversee its Section 1557 program).

<sup>94</sup> See, e.g., OCR ENFORCEMENT OF OAACS, *supra* note 90, at 4 (requiring the state office to conduct follow up assessments on its newly developed race-neutral policies).

<sup>95</sup> See STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 3 (“When possible, that quantitative data should be audited by an agency or institution independent from actors within the system to avoid biased statistical reporting.”).

<sup>96</sup> See *id.* at 3 (“If the use of a particular pretrial risk assessment instrument by itself does not result in an independently audited, measurable decrease in the number of people detained pretrial, the tool should be pulled from use until it is recalibrated to cause demonstrably decarceral results.”). While some scholars speculate that attempts to remedy disparate impact by intentionally rebalancing AI in favor of groups that have historically been marginalized will run afoul the

entities commitment to nondiscrimination, “voluntary action” is also the preferred remedy under Section 1557 regulations.<sup>97</sup>

To ensure that these nondiscrimination policies are in effect on the ground, providers (e.g., physicians, nurses) and administrative staff (e.g., medical billers, intake and discharge staff, and even personnel working in the social services offices at hospitals and clinics) need to be trained on Section 1557.<sup>98</sup> As of now, AI in healthcare does not have primacy—human beings are still able and expected to review the AI’s predictions and decide how to proceed. Covered entities need to clarify what conduct is not permitted (e.g., accepting AI determinations without additional review) and consistently discipline staff who violate the AI-nondiscrimination policies.<sup>99</sup> This human aspect should be more than a last-line-of-defense to detect medical or coding errors, but should also recognize discrimination under Section 1557. For example, if an AI makes a prediction based on a patients’ answer to a questionnaire, then the medical staff should recognize that such prediction might run afoul Section 1557 if the patient did not receive interpreter services if the patient is LEP.

A hallmark of a robust compliance program is an effective reporting and complaint mechanism. Covered entities should already have anonymous reporting

---

Supreme Court’s decision in *Ricci v. Stefano*, others argue that fixing AI to avoid foreseeable disparate impact liability is permissible. Compare Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 694–95 (2017) (“The holding in *Ricci* suggests that we cannot solely rely on auditing for legal reasons . . . If an agency runs an algorithm that has a disparate impact, correcting those results after the fact will trigger the same kind of analysis as New Haven’s rejection of its firefighter test results.”), with Mark MacCarthy, *Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms*, 48 CUMB. L. REV. 67, 130 (2017–18) (arguing that *Ricci* does not compel such strict barriers and as long as an agency can foresee disparate impact through “a strong belief in evidence” it can adopt a less biased policy if it does not require taking a benefit away from the individuals who received it under the original policy).

<sup>97</sup> 45 C.F.R. § 92.6(b); see, e.g., OCR ENFORCEMENT OF OAACS, *supra* note 90, at 1 (explaining that the state office voluntarily developed its own “Disproportionality Action Plan” to end the practice of referring and transferring children based on their race, prior to an OCR investigation). Additionally, it would be wise to involve corporate counsel in the ongoing monitoring of healthcare AI systems. See STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 4–5 (“Defense counsel must be included in the process of selecting, calibrating, designing, shaping, and testing a pretrial risk assessment instrument and included in the ongoing evaluation of the tool.”).

<sup>98</sup> See, e.g., OCR Enforcement of Touro Infirmary, *supra* note 87 (requiring Touro to train emergency department staff to identify domestic violence victims); OCR ENFORCEMENT OF PDHS, *supra* note 88, at 13 (requiring the Pennsylvania agency to train all relevant staff who have regular contact with benefits applicants on interpreter-services policies); OCR ENFORCEMENT OF MMI, *supra* note 89, at 11 (requiring the Institute to train all staff about Section 1557 and policies regarding provision of auxiliary aids); OCR ENFORCEMENT OF OAACS, *supra* note 90, at 6 (requiring the state office to train all staff on Title VI and internal nondiscrimination obligations).

<sup>99</sup> See, e.g., OCR ENFORCEMENT OF PDHS, *supra* note 88, at 9 (prohibiting employees from using family or friends of patients as interpreters).

lines and enforced anti-retaliation policies (for internal whistleblowers);<sup>100</sup> it should be clear to employees, senior officers, and vendors that these mechanisms can also receive Section 1557 complaints. Similarly, covered entities should implement grievance procedures (for patients) to report possible Section 1557 violations when AI is used to make treatment decisions, billing activities, etc. relating to a patient.<sup>101</sup> One can question the usefulness of reporting mechanisms when most of an AI's decision-making is made behind a digital curtain—will patients even know if they've been discriminated against? In order to make grievance procedures for patients meaningful, covered entities should (a) notify patients when AI is used in relation to them and (b) explicitly advertise the grievance procedures where patients are likely to see them.<sup>102</sup>

A criticism of implementing compliance policies around Section 1557-AI claims is that covered entities can use their robust policy as a shield in OCR investigations. Given that healthcare AI is still (relatively) in its infancy and that OCR has not yet addressed it, OCR may defer to covered entities' interpretation of appropriate Section 1557 compliance in this realm. This can lead to the obvious pitfall of OCR never engaging in meaningful review of discrimination when AI is used. One method of combatting this is of course for the healthcare industry to set a high standard for itself; covered entities should properly vet the AI developers they partner with to ensure they incorporate Section 1557 principles in their original designs.<sup>103</sup>

From the perspective of a covered entity, Section 1557 is an important legal obligation for one major reason: failure to comply *can* lead to termination of federal financial assistance.<sup>104</sup> In most cases, loss of such assistance is fatal to the continued existence of the covered entity.<sup>105</sup> However, the threat of Section 1557 enforcement

---

<sup>100</sup> 45 C.F.R. § 92.7(b) (requiring covered entities with more than fifteen employees to adopt Section 1557 grievance procedures).

<sup>101</sup> See, e.g., OCR ENFORCEMENT OF OAACS, *supra* note 90, at 6 (requiring the development of a complaint resolution procedure); OCR ENFORCEMENT OF MMI, *supra* note 89, at 6 (requiring the Institute to revise its grievance procedures); OCR ENFORCEMENT OF PDHS, *supra* note 88, at 7 (requiring the state agency to update its standard procedures for receiving limited-English-proficient complaints).

<sup>102</sup> See, e.g., OCR ENFORCEMENT OF PDHS, *supra* note 88, at 10 (requiring PDHS to use posters, a website, written notices, and telephonic interpreters at points of contact to ensure patients with LEP needs are identified).

<sup>103</sup> Such vetting indicates a need for regulation at the *design* state of healthcare AI, not just on the *application* stage that OCR oversees. This will be discussed further in Part IV.

<sup>104</sup> 45 C.F.R. § 92.302(c).

<sup>105</sup> For a Chicago hospital that lost its federal money due to investigations about abuse, 80% of their patients were insured Medicare or Medicaid; its third highest source of revenue was from BlueCross, under which only 9% of patients were insured. See Duaa Eldeib, *Chicago Psychiatric Hospital Will Lose Federal Money, and Its License Is Threatened After Allegations of Abuse*, PROPUBLICA ILLINOIS (Dec. 27, 2019, 7:22 AM), <https://www.propublica.org/article/chicago->



lacks the oomph that other healthcare enforcement actions carry. When OCR investigations result in Voluntary Resolution Agreements, if the party fails to substantially comply with the Agreement, the result is that the “parties will confer and attempt to reach agreement as to what steps may be necessary to resolve the compliance issues to both parties’ satisfaction.”<sup>106</sup> Compare this with enforcement threats in other areas of healthcare—such as fraud and abuse laws—where the failure to comply with the law is met with hefty fines.<sup>107</sup>

Moreover, OCR’s enforcement of Section 1557 has been relatively sparse and limited in scope—particularly under the Trump Administration.<sup>108</sup> In mid-2019, the Department of Health and Human Services proposed “substantial revisions” to Section 1557—namely eliminating protection of transgender individuals, adopting blanket religious freedom exemptions for providers, and weakening protections for individuals with limited English proficiency.<sup>109</sup> Conscience and religious freedom were the main focal points of OCR under the Trump Administration, beginning with Trump’s Executive Order “Promoting Free Speech and Religious Liberty”<sup>110</sup> and the creation of the Conscience and Religious Freedom Division within OCR.<sup>111</sup> By their own report, enforcing civil rights was only about a quarter of OCR’s workload, and the majority of that time has focused

---

psychiatric-hospital-will-lose-federal-money-and-its-license-is-threatened-after-allegations-of-abuse (noting that patients are also harmed when hospitals lose federal funding).

<sup>106</sup> See, e.g., OCR ENFORCEMENT OF MMI, *supra* note 89, at 4.

<sup>107</sup> False Claims Act (FCA), 31 U.S.C. § 3729(a)(1)(G) (2018) (\$5,000–\$10,000 per claim); Anti-Kickback Statute, 42 U.S.C. § 1320a-7b(b) (2018) (\$100,000); Civil Monetary Penalties Law, 42 U.S.C. § 1320a-7a(o) (2018) (\$50,000 per false claim); 42 U.S.C. § 1395nn(g)(3) (2018) (\$15,000 per service).

<sup>108</sup> Letter from Leadership Conference on Civil and Human Rights to Thomas E. Price, Sec’y, U.S. Dep’t of Health & Human Servs. & Roger Severino, Director, Office for Civil Rights, U.S. Dep’t of Health & Human Servs. (Mar. 30, 2017), [http://civilrightsdocs.info/pdf/policy/letters/2017/OCR%20Letter\\_1557\\_3-30-17.pdf](http://civilrightsdocs.info/pdf/policy/letters/2017/OCR%20Letter_1557_3-30-17.pdf) (noting that under the Obama Administration, OCR “accepted, investigated, and resolved discrimination complaints alleged violations under Section 1557”) [hereinafter Leadership Conference Letter].

<sup>109</sup> See Nondiscrimination in Health and Health Education Programs or Activities, 84 Fed. Reg. 27,846 (June 14, 2019).

<sup>110</sup> Exec. Order No. 13,798, 82 Fed. Reg. 21,675 (May 4, 2017).

<sup>111</sup> See OFFICE FOR CIVIL RIGHTS, *Conscience and Religious Freedom News Releases & Bulletins*, HHS.GOV, <https://www.hhs.gov/conscience/newsroom/index.html> (last visited Apr. 17, 2021); see also Emmarie Huetteman, *At Trump administration’s new health office, ‘civil rights’ means doctors have a right to say no to patients*, FIERCE HEALTHCARE (Mar. 6, 2018, 11:00 AM), <https://www.fiercehealthcare.com/practices/civil-rights-means-doctors-right-to-say-no-to-patients> (noting that an HHS spokesperson reported a “clear surge” in conscience-related complaints under the Trump Administration).

on the conscience rights rule.<sup>112</sup> Enforcement of Section 1557 is necessary because it can prevent and remedy discrimination at “every step in the healthcare system, from obtaining insurance coverage to receiving a proper diagnosis and treatment.”<sup>113</sup> The limited nature of OCR enforcement of Section 1557 generally may lead patients to feel unprotected by the agency when AI is used.

### *B. Section 1557: Private Right of Action*

Healthcare consumers themselves have a private right of action under Section 1557 for allegations of *intentional discrimination*, giving them the ability to file a lawsuit against a covered entity directly rather than filing a complaint with OCR. To state a claim under Section 1557, plaintiffs must show that they (1) were a member of a protected class, (2) qualified for the benefit/program at issue, (3) suffered an adverse action, and (4) the adverse action gave rise to an inference of discrimination.<sup>114</sup>

However, there is a circuit split as to whether Section 1557 provides for a private right of action for *disparate-impact* claims. Some courts have permitted disparate-impact liability for all grounds prohibited under Section 1557; they justify that it would lead to absurd and inconsistent results if only some plaintiffs could proceed without a showing of intentional discrimination.<sup>115</sup> Other courts, however, have held that Section 1557 allows for disparate impact liability only where the originating statutory basis provided for it.<sup>116</sup> Recall that Section 1557 incorporates different civil rights statutes (e.g., Civil Rights Act), not all of which allow for actions under a disparate-impact theory.<sup>117</sup> Despite the fact the OCR issued a guidance letter explicitly stating that the agency recognized a private right of action for disparate impact discrimination, these courts argue that the letter is clearly

---

<sup>112</sup> DEP’T OF HEALTH & HUMAN SERVS., FISCAL YEAR 2021 BUDGET IN BRIEF 171–72 (2021), <https://www.hhs.gov/about/budget/index.html> (noting that 73% of the workload is on HIPAA).

<sup>113</sup> Leadership Conference Letter, *supra* note 108, at 1.

<sup>114</sup> *See Griffin v. General Electric Co.*, 752 Fed. Appx. 947, 949 (11th Cir. 2019).

<sup>115</sup> *Rumble v. Fairview Health Services*, No. 14-cv-20137 (SRN/FLN), 2015 WL 1197415, at \*11–12 (D. Minn. Mar. 16, 2015) (holding that Congress referenced the four civil rights statute merely to identify the grounds for prohibited discrimination, but reaching the question of what standard does apply).

<sup>116</sup> *Briscoe v. Health Care Services Corp.*, 281 F. Supp.3d 725, 738 (N.D. Ill. 2017) (sex discrimination under Title XI of the Civil Rights Act); *Southern Pennsylvania Transport Authority v. Gilead Sciences, Inc.*, 102 F. Supp. 3d 688 (E.D. Penn. 2015) (race discrimination under Title VI of the Civil Rights Act).

<sup>117</sup> There is no private right of action for disparate impact claims under Title VI of the Civil Rights Act. *See Alexander v. Sandoval*, 532 U.S. 275 (2001). There is, however, a private right of action under Title IX, the Age Discrimination Employment Act, and the Rehabilitation Act of 1973. *See Gesber v. Lago Vista Independent School District*, 524 U.S. 274 (1998); *Smith v. City of Jackson*, 544 U.S. 228 (2005); *Alexander v. Choate*, 469 U.S. 287 (1985).

contrary to Congress’s intent and does not deserve *Chevron* deference.<sup>118</sup> Judge Sutton’s commentary in *Doe v. BlueCross BlueShield of Tennessee* sums up the mindset of these courts: “Treating similarly situated people differently goes to the heart of invidious discrimination. But treating differently situated people differently usually counts as equal justice under law.”<sup>119</sup>

Due to the circuit split on disparate-impact liability for Section 1557, claims can be fatal for many individuals seeking relief through a private lawsuit for discrimination they faced because of healthcare AI. As the UnitedHealth case study reported, factors related to a suspect class (in that case, race) are not expected to be explicitly included in the AI’s algorithm; yet biases can enter and manifest through *other* factors (in that case, cost history).<sup>120</sup> The types of claims that private plaintiffs might pursue will clearly fit the definition of disparate-impact liability—i.e., facially neutral policies that have a discriminatory effect. In the circuits where disparate-impact liability is not always permitted, patients seeking redress for the effects of AI biases on racial grounds are unlikely to prevail.<sup>121</sup>

Even in the states that do permit disparate-impact liability—or in the rare cases where a patient can show intentional discrimination—the odds of plaintiffs prevailing are low. Section 1557 lawsuits are few and far in between.<sup>122</sup> When brought, they are often decided in favor of the healthcare covered entity.<sup>123</sup> As a result, patients who feel the brunt of bias in healthcare AI have little hope for recourse under Section 1557 complaints to OCR or private lawsuits.

---

<sup>118</sup> *E.g.*, *Briscoe v. Health Care Services Corp.*, 281 F. Supp.3d 725, 738 (N.D. Ill. 2017) (noting that if Congress intended for there to be a single standard of liability, it would not have included each enforcement mechanism of each statute it incorporated); *see also Rumble*, 2015 WL 1197415, at \*10 (holding that OCR’s opinion letter is not controlling, but could receive *Skidmore* deference as a persuasive document).

<sup>119</sup> *Doe v. BlueCross BlueShield of Tenn., Inc.*, 926 F.3d 235, 237 (6th Cir. 2019).

<sup>120</sup> *Obermeyer et al.*, *supra* note 1, at 452.

<sup>121</sup> *See Doe*, 926 F.3d at 241 (holding that a policy that required all patients to switch to a special network pharmacy for high-cost drugs, rather than their local pharmacy, is neutral on its face and did not intentionally discriminate against patients with HIV on the basis of disability).

<sup>122</sup> At the time of publication, the author could only find 19 pending or decided cases referencing Section 1557 when running the appropriate Westlaw search. (search terms “42 U.S.C. 18116,” then Citing References and “Yes” to “Referenced in Notes of Decisions”).

<sup>123</sup> *See, e.g.*, *Tovar v. Essentia Health*, 857 F.3d 771, 781 (8th Cir. 2017) (Benon, J., concurring in part) (dismissing because the plaintiff named the third-party administrator of a self-insured plan, and not the employer, as the defendant); *see also Griffin v. Verizon Commc’ns, Inc.*, 746 Fed. App’x 873, 874 (11th Cir. 2018) (finding that the alleged plaintiff could not demonstrate discrimination by an employer-based health plan based on its “litigation conduct” when it chose not to assert a defense in actions by white, male providers but asserted it in the case of a black, female provider).

## V. COMPLIANCE CHALLENGES

### A. Licensing

Federal agencies, such as the FDA or OCR in Health and Human Services, do not have authority to regulate healthcare providers themselves—that task is left to the states. Various state medical licensing boards regulate the practice of medicine by physicians, as well as nurses and technicians. While the Federation of State Medical Boards similarly recognizes the need for regulation around AI,<sup>124</sup> state licensing boards may not be able to regulate AI because it fails to meet the most basic hook for their authority—practice of medicine.<sup>125</sup> As of now, physicians maintain primacy; it is presumed that the AI does not *make* decisions, it simply recommends them.<sup>126</sup>

Licensed physicians are held to certain standards of conduct based on ethical guidelines (transparency, truthfulness) and legal obligations (reasonable care, confidentiality, informed consent).<sup>127</sup> Nearly all medical professionals are required to participate in continuing education (CE) hours as well in order to maintain their license; some states require certain professionals to take CE hours related to a subject of cultural competency (e.g., LGBTQ, HIV/AIDS).<sup>128</sup> Federal guidelines fill in gaps where states have not developed such standards. For example, OCR’s *Title VI Prohibition Against National Origin Discrimination Affecting Limited English Proficient Persons*,<sup>129</sup> or the Office of Minority Health’s *National Standard for Culturally and Linguistically Appropriate Services (CLAS) in Health*

---

<sup>124</sup> See *Artificial Intelligence*, FED’N OF ST. MED. BOARDS, <https://www.fsmb.org/ai/> (last visited Mar. 5, 2020). The FSMB has also created a workgroup consisting of stakeholders, including from the American Medical Association, to provide resources to state medical boards on the effects of AI on patient safety and quality of healthcare. See FED’N OF ST. MED. BOARDS, Res. 18–6, WORKGROUP ON AI AND ITS POTENTIAL IMPACT ON PATIENT SAFETY AND QUALITY OF CARE IN MEDICAL PRACTICE (2018), <https://www.fsmb.org/siteassets/artificial-intelligence/pdfs/2018-ai-resolution.pdf>.

<sup>125</sup> See Terry, *supra* note 9, at 154.

<sup>126</sup> See generally U.S. FOOD & DRUG ADMIN., CLINICAL DECISION SUPPORT SOFTWARE DRAFT GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF (2019), <https://www.fda.gov/media/109618/download> (discussing the support role of CDS, including machine learning, but also discussing SaMD that drives clinical management and SaMD that treats, prevents, mitigates, diagnoses, screens, and detects conditions and diseases).

<sup>127</sup> *Id.*

<sup>128</sup> *Board of Medicine*, DC.GOV, <https://dchealth.dc.gov/bomed> (last visited Mar. 10, 2020) (requiring LGBTQ and HIV/AIDS cultural competency courses for Medical Doctors, Doctors of Osteopathic Medicine, and Physicians Assistants).

<sup>129</sup> U.S. DEP’T OF HEALTH & HUMAN SERVS., GUIDANCE TO FEDERAL FINANCIAL ASSISTANCE RECIPIENTS REGARDING TITLE VI PROHIBITION AGAINST NATIONAL ORIGIN DISCRIMINATION AFFECTING LIMITED ENGLISH PROFICIENT PERSONS (2004).

*and Health Care*.<sup>130</sup> While neither comprehensive nor necessarily binding, the guidance and CE requirements still ensure that there is at least *some* understanding of health disparities in certain communities. There is no guarantee that such guidelines are applied consistently or uniformly, but there is something to be said about human element behind traditional healthcare.<sup>131</sup>

In the context of pretrial detainment decisions, advocacy groups recommend that courts abide by an AI determinations only if it recommends release (and if the algorithm does not recommend immediate release, it should recommend a pretrial hearing involving rigorous safeguards).<sup>132</sup> Borrowing from this approach, state licensing boards can write their policies to ensure that that if AI is used, it is used only in contexts with the lowest potential for harm (e.g., for treatment maintenance, but not diagnosis; or to diagnose common colds, but not cancers). In line with this least-approach harm, providers should include in their compliance program that providers can follow diagnostic AI determinations if it *does* make a diagnosis, but if the AI *does not* make a diagnosis there should be a rigorous independent review by the physician to protect against delayed diagnosis.<sup>133</sup>

### *B. Malpractice Liability*

Traditional tort liability for medical malpractice also extends to discrimination by healthcare providers. There are mountains of evidence—both statistical and anecdotal—of discrimination against patients based on their race, gender, disability and other characteristics. For example, physicians overestimate pain tolerance of patients of color, leading to systemic undertreatment.<sup>134</sup> Female patients similarly have their symptoms downplayed by physicians, and patients belonging to more than one marginalized group experience compounding discrimination.<sup>135</sup> Providers can be held liable for providing inferior treatment,

---

<sup>130</sup> THINKCULTURALHEALTH, *National Culturally and Linguistically Appropriate Services Standards*, HHS.GOV, <https://thinkculturalhealth.hhs.gov/clas/standards> (last visited Apr. 8, 2021).

<sup>131</sup> A physician who is also from an underrepresented community may be more attune to such issues. A nurse practitioner that who has one patient who is limited English proficient may be more likely to seek the appropriate services with similarly situated patients in the future.

<sup>132</sup> See STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 4.

<sup>133</sup> See *id.* at 4.

<sup>134</sup> See Hoffman et al., *supra* note 7, at 4296; see also P.R. Lockhart, *What Serena Williams's scary childbirth story says about medical treatment of black women*, VOX (Jan. 11, 2018, 4:40 PM), <https://www.vox.com/identities/2018/1/11/16879984/serena-williams-childbirth-scary-black-women> (reporting that black women are three to four times more likely to die from maternity complications).

<sup>135</sup> See Katie Klabusich, *It's Not All in Your Head: New Book Sheds Light on 'Bad Medicine and Lazy Science' Harming Women*, REWIRE NEWS GROUP (Mar. 7, 2018, 10:58 AM), <https://rewire.news/article/2018/03/07/not-head-new-book-sheds-light-bad-medicine-lazy-science->

misdiagnosing, or denying care to patients based on their protected characteristics—whether intentional or unintentional.<sup>136</sup> The existence of tort liability for discriminatory treatment by providers raises the question: Is discrimination by healthcare AI a cause of action for medical practice?

Scholars and healthcare providers alike recognize that the use of AI in healthcare, particularly for clinical decision-making, does not amount to “substitution” of the provider’s role.<sup>137</sup> Regardless of a physician’s reliance on AI, the responsibility of clinical decision-making lies with the physician: The AI does not *practice* medicine.<sup>138</sup>

Physicians can be subject to tort liability when their diagnosis and treatment decisions fall below the standard of the customary practice.<sup>139</sup> Where any new practices are used, it is naturally difficult for the physician to establish that she comported with “custom.”<sup>140</sup> When AI is used and results in injury to a patient (e.g., from misdiagnosis), particularly injury arising from a bias, is the physician liable for the AI’s decision? To date, no court has applied the physician liability standard where AI was at issue.<sup>141</sup> Similar to the challenges created by lack of regulation, lack of tort precedent around healthcare AI creates a compliance challenge for providers.

Malpractice litigation around healthcare AI may center on the established safety and quality of the technology. Nicholas Terry points out that in many domains, automated technology is expected or even preferred (e.g., automation in

---

that-harm-women/ (noting that women of color and low-income women face discrimination at high rates).

<sup>136</sup> See *Can discrimination be a form of medical malpractice?*, CRANWELL & MOORE P.L.C., ATTORNEYS AT LAW (Apr. 27, 2018), <https://www.cranwellmoorelaw.com/blog/2018/04/can-discrimination-be-a-form-of-medical-malpractice/> (explaining that preventable medical errors arising out of a physician’s prejudice are actionable). However, not all disparate treatments in the medical profession are discriminatory. See generally *Butler v. Flint Goodrich Hosp. of Dillard*, 607 So.2d 517, 521 (La. 1992) (holding that a statutory cap on malpractice judgments, which essentially treats people with more costly injuries differently than those with less costly injuries, was not unconstitutional).

<sup>137</sup> See Terry, *supra* note 9, at 148. Although beyond the scope of this Article, Terry further notes that patients in underdeveloped nations may welcome AI as the primary provider—not as a substitute, but as “otherwise unobtainable healthcare.”

<sup>138</sup> See *id.* at 148. Babylon Health, for example, states on their “About” page: “Babylon’s AI services provide health information only, and do not provide diagnosis.” BABYLON HEALTH, <https://www.babylonhealth.com/ai> (last visited Mar. 5, 2020).

<sup>139</sup> See Ben A. Rich, *Medical Custom and Medical Ethics: Rethinking the Standard of Care*, 14 CAMBRIDGE Q. HEALTHCARE ETHICS 27, 28 (2005).

<sup>140</sup> See Zach Harned et al., *Machine Vision, Medical AI, and Malpractice*, HARV. J. L. & TECH. DIGEST 6 (2019), <https://jolt.law.harvard.edu/assets/digestImages/PDFs/Harned19-03.pdf>.

<sup>141</sup> See generally *id.* at 6.

commercial airlines).<sup>142</sup> However, the FDA has issued guidance (read *warnings*) stating that the “safety and effectiveness” of certain robotic devices—such as robotically-assisted surgical devices in mastectomy and cancer treatments—has not been established.<sup>143</sup> Where AI technology is untested (or has been shown to have negative biases), providers should avoid using them in their clinical decisions or consult with the relevant agencies for advisory opinions. Healthcare entities should engage in regular and thorough review to ensure the technology they opt for is not later revealed to have discriminatory effects.

Moreover, physicians are still expected to have ultimate review of their AI’s clinical determinations. Regardless of advancements in AI, it is unlikely that courts will permit patients to bring malpractice lawsuits against the AI system. Therefore, malpractice litigation may apply scope-of-practice doctrines to AI, such as by allowing AI to make decisions under a physician’s “standing orders” (i.e., an algorithm).<sup>144</sup> Another analysis courts might use is to treat physicians’ use of AI as legitimate consultation with an expert.<sup>145</sup> For example, when AI is used to make a radiology decision under this approach, the physician must establish that the AI is properly “trained,” used a patient’s specific clinical history and findings to reach its decision, identified similar cases, and can “point” to the relevant parts of a patient’s scan.<sup>146</sup> Ultimately, AI may advance to the point that using AI is customary practice and drawing analogies to expert-consultations is unnecessary.<sup>147</sup> An additional compliance challenge may arise from the fact that there are multiple entities in the distribution chain of healthcare AI, including the manufacturer that creates it (products liability), to the hospitals and clinics that

---

<sup>142</sup> See Terry, *supra* note 9, at 174.

<sup>143</sup> *Caution When Using Robotically-Assisted Surgical Devices in Women’s Health including Mastectomy and Other Cancer-Related Surgeries: FDA Safety Communication*, U.S. FOOD & DRUG ADMIN. (Feb. 28, 2019), <https://www.fda.gov/medical-devices/safety-communications/caution-when-using-robotically-assisted-surgical-devices-womens-health-including-mastectomy-and>.

<sup>144</sup> See *supra* Part IV(A) Licensing.

<sup>145</sup> See Harned et al., *supra* note 140, at 6–7.

<sup>146</sup> See *id.* at 7.

<sup>147</sup> This point raises an interesting question in terms of the future of malpractice liability: Will AI alter our interpretation of the duty of care that physicians owe their patients? Currently, the standard of care for malpractice liability relies on expert testimony of what other physicians would do in the same or similar circumstances—this question necessarily includes a physician’s ability to empathize with and appreciate a patient’s circumstances. See *id.* at 7. As physicians increasingly rely on AI decisions, this element of human empathy may gradually be removed from the standard of care determinations.

adopt it (facilities and equipment), and to the physicians that use and review its decisions (malpractice).<sup>148</sup>

Tort liability presents many risks for providers, so their compliance programs seek to mitigate such risks (e.g., through reporting mechanisms). Given the numerous open questions presented in this section, healthcare providers should, again, take a cautious approach. Since the law treats providers as having primacy, they should always thoroughly review all AI determinations—whether it is a new decision (e.g., an initial diagnosis, or new treatment plan) or a decision to make *no change* to the patient’s treatment, diagnosis, or prescription. Reviewing even these seemingly low-risk decisions is important from the provider’s perspective because even a decision to do nothing is still a decision for which the physician can be liable.

## VI. RECOMMENDATIONS

Thus far, this Article has discussed what is lacking in terms of antidiscrimination in healthcare AI. The FDA’s enforcement is only emerging and does not fully capture the potential for discriminatory effects by AI. Section 1557, while binding on covered entities, has taken a backseat to OCR’s other enforcement efforts, and private rights of action are largely unavailable to patients seeking disparate-impact liability. Additionally, OCR has been silent on the applicability of Section 1557 on the use of AI tools. And state licensing boards lack the statutory hook to govern the way healthcare providers rely on AI. A new approach is necessary to ensure that nondiscrimination is a valued legal principle in healthcare AI, two questions remain: What is the right approach? And who will enforce it?

### A. Industry Standards and Internal Compliance

It would be wise to require nondiscrimination at the start—at AI development. One method would be for the AI industry to adopt its own standards that reflect nondiscrimination. In the absence of legal obligations, some scholars suggest private companies adopt internal standards to reflect an ethical approach to creating AI.<sup>149</sup> This can take the form of internal standards (e.g., Facebook, IBM, Microsoft) or standards set forth by professional associations (e.g., Association for the Advancement of Artificial Intelligence, Association of Computing Machinery, Institute of Electrical and Electronics Engineers).<sup>150</sup> Some organizations, like the Institute of Electrical and Electronics Engineers’ 2016 report, “Ethically Aligned Design,” emphasized “human norms and values” and “value sensitive or value-

---

<sup>148</sup> *Taylor v. Intuitive Surgical, Inc.*, 389 P.3d 517 (Wash. 2017) (addressing the question of which members of the distribution chain are subject to liability and holding that product liability under the Washington Consumer Protection Act requires the device manufacturer to provide product warnings to the hospital *and the physician*).

<sup>149</sup> See Katyal, *supra* note 22, at 108–09.

<sup>150</sup> See *id.* at 109.



based design.”<sup>151</sup> The British Computer Society’s code of conduct requires individuals must “promote equal access to the benefit of IT and promote the inclusion of all sectors of society whenever opportunities arise.”<sup>152</sup> These principles can (and should) be adopted by trade organizations of AI developers and should be incorporated into the codes of conduct, policies and procedures, and human impact statements of individual AI companies.

Google, for example, lists its principles and objectives for its AI products publicly. One such objective is to “avoid creating or reinforcing unfair bias.”<sup>153</sup>

AI algorithms and datasets can reflect, reinforce, or reduce unfair biases. We recognize that distinguishing fair from unfair biases is not always simple and differs across cultures and societies. We will seek to avoid unjust impacts on people, particularly those related to sensitive characteristics such as race, ethnicity, gender, nationality, income, sexual orientation, ability, and political or religious belief.<sup>154</sup>

IBM similarly includes “fairness” as an element in its “Five Areas of Ethical Focus.”

AI provides deeper insight into our personal lives when interacting with our sensitive data. As humans are inherently vulnerable to biases, and are responsible for building AI, there are chances for human bias to be embedded in the systems we create. It is the role of a responsible team to minimize algorithmic bias through ongoing research and data collection which is representative of a diverse population.<sup>155</sup>

Since not all bias can be eliminated from AI decision-making (despite the data-vetting methods discussed in Part III(A)), healthcare entities must consider methods to counterbalance bias in their compliance programs. Under industry standards, manufacturer’s may already be expected to engage in independent

---

<sup>151</sup> INSTITUTE FOR ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE), ETHICALLY ALIGNED DESIGN 29, 39 (2016), [http://standards.ieee.org/develop/indconn/ec/ead\\_v1.pdf](http://standards.ieee.org/develop/indconn/ec/ead_v1.pdf)

<sup>152</sup> *BCS Code of Conduct*, BRITISH COMPUT. SOC’Y, <https://www.bcs.org/category/6030> (last visited Mar. 26, 2020).

<sup>153</sup> Sundar Pichai, *AI at Google: our principles*, GOOGLE (June 7, 2018), <https://www.blog.google/technology/ai/ai-principles>.

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

<sup>155</sup> IBM, EVERYDAY ETHICS FOR ARTIFICIAL INTELLIGENCE 32, 38–39 (2019), <https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf>. IBM further makes several recommendations: (1) investigating and understanding intentional and unintentional biases; (2) design and develop without intentional biases and schedule team reviews for unintentional biases, including stereotyping, confirmation biases, and “sunk cost” bias; and (3) include a “feedback mechanism or open dialogue” with users to allow user-identified biases.

review of their products for validity and quality;<sup>156</sup> such standards can be expanded to require independent review for biases. Key to achieving the nondiscrimination principles discussed here will be transparency—entities should develop their own factors for ensuring that information about their AI technology is readily available to patients, providers, and regulators. This information may include (1) a description of the AI’s design and testing, (2) the elements on which the AI makes predictions, the weights assigned to those elements, and (3) the source of the training dataset and the “outcome data” used to validate the AI.<sup>157</sup> This level of transparency will not only assist regulators to better understand the landscape of healthcare AI and develop future regulations, but it also provides patients with a meaningful opportunity to challenge the use of AI if they experience harmful effects.

While the opportunity to challenge is highly recommended as a counterbalance method to AI bias in other industries (e.g., criminal justice),<sup>158</sup> the opportunity to file for a hearing will not help a patient who has been incorrectly diagnosed by a physician or whose coverage has been denied by an insurer.<sup>159</sup> Patients (and their lawyers) are not versed in medicine, they do not possess the necessary expertise to dispute the correctness of medical diagnosis or appropriateness of a treatment plan, particularly to effectively argue that AI bias played a role in those decisions. Healthcare entities should consider methods to make the AI open to a meaningful challenge, such as implementing a panel where a harmed patient and her counsel can access professional medical opinions to build their case.<sup>160</sup> Additionally, covered entities should consider other methods to counterbalance the effects of bias in healthcare AI when it is used—such as a rebuttable presumption against the AI decision or *de novo* review of its decisions when a patient challenges an AI determination.

---

<sup>156</sup> SAMD GUIDANCE, *supra* note 44, at 16. These SaMD guidelines were created by IMDFR, then adopted by FDA.

<sup>157</sup> STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 7.

<sup>158</sup> *See id.* at 7.

<sup>159</sup> Fatma E. Marouf, *Alternatives to Immigration Detention*, 38 CARDOZO L. REV. 2141, 2144–45 (2017). Even the “open to challenge” option can be problematic: A study of ICE found that the agency overrode its algorithm-based risk-assessment tool’s recommendations in 19% of cases where LGBT people were involved. In developing compliance programs, covered entities should consider the risk of confirming (or injecting) bias if they offer appeals options.

<sup>160</sup> *See* STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 6. Additionally, healthcare entities should ensure that the patient’s counsel has adequate time to meet with the patient and learn their circumstances (i.e., factors *not* captured in the AI determination) prior to any administrative hearing or other challenge procedure—this may involve extended time to file appeals or requests for reconsideration.

### B. Recommended Regulations and a “Super Regulation”

Some argue that regulators, not private companies, should take the lead in developing policies around AI-informed tools.<sup>161</sup> The findings in the UnitedHealth case study should indicate to the AI industry that companies need to be robust in their nondiscrimination assessments—merely cutting out factors such as race or disability is insufficient to eliminate disparate-impact to those characteristics.<sup>162</sup> Professor Sean K. Hallisey suggests an “AI Data Transparency Model” that shifts the focus from regulating the AI algorithms and developers towards regulating the data itself through auditing and certification requirements.<sup>163</sup> The FDA can adopt guidelines (or add guidelines to its proposed frameworks) that reflect nondiscrimination standards in the SaMD and CDS which it approves. However, the descriptor “medical device”—the hook for FDA regulation—fails to fully capture the cognitive-like characterization of AI.<sup>164</sup> As discussed earlier, the FDA’s limited authority would still leave healthcare AI that is used for billing or other administrative purposes unregulated.

Another approach is to regulate the usage of AI, as opposed to its creation. There are some bases to find that state licensing boards have authority to regulate healthcare AI: (1) future AI, particularly diagnostic and procedural ones, may actually be able to *practice* medicine and (2) other doctrines permit licensure regulations, particularly the corporate practice of medicine and scope of practice (SOP).<sup>165</sup> SOP guidelines for nurse practitioners who, depending on jurisdiction, can diagnose and treat with or without physician involvement can be used as a model for SOP around healthcare AI.<sup>166</sup> The application of SOP doctrine leads to

---

<sup>161</sup> Hayley Tsukayama & Jamie Williams, *If A Pre-Trial Risk Assessment Tool Does Not Satisfy These Criteria, It Needs to Stay Out of the Courtroom*, ELEC. FRONTIER FOUND. (Nov. 6, 2018), <https://www.eff.org/deeplinks/2018/11/if-pre-trial-risk-assessment-tool-does-not-satisfy-these-criteria-it-needs-stay>. While Tsukayama and Williams’s argument was in the context of the criminal justice system, the principles they discuss may translate to the healthcare industry.

<sup>162</sup> See *id.* Removing just the suspect factors (e.g., race, gender) will not remedy the inherent biases in the data—a concept known as “omitted variable bias.” See also Heidi Ledford, *Millions of black people affected by racial bias in health-care algorithms*, NATURE (Oct. 26, 2019), <https://www.nature.com/articles/d41586-019-03228-6#ref-CR1>. The problem is further exacerbated by the lack of diversity among software developers because they are less likely to anticipate potential biases or consider including certain mitigating data-elements.

<sup>163</sup> Schlomit Yanisky-Ravid & Sean K. Hallisey, “Equality and Privacy by Design”: A New Model of Artificial Intelligence Data Transparency Via Auditing, Certification, and Safe Harbor Regimes, 46 FORDHAM URB. L.J. 428, 434 (2019).

<sup>164</sup> Terry, *supra* note 9, at 172.

<sup>165</sup> See *id.* at 154–55. The CMP doctrine prohibits entities that cannot be personally licensed (e.g., corporations) from either practicing medicine or employing physicians to practice (with the exception of hospitals). Nicholas Terry hypothesizes that AI at some point may advance to the point that it *cannot* leverage the licensed-hospital exception.

<sup>166</sup> See *id.* at 155.

an interesting questions: If an AI is “licensable” to what extent can it practice and be developed for one purpose and be tasked for another?

Finally, some scholars suggest an entirely new regulatory regime—perhaps even the creation of a new agency—to oversee development of all AI technology (not just healthcare) through a “super regulator” or a third-party auditing mechanism.<sup>167</sup> “There are ways to minimize bias and unfairness in pretrial risk assessment, but it requires proper guidance and oversight.”<sup>168</sup> California Senate Bill 10 is an example of an effort to regulate AI that fell short. The Bill sought to implement AI for pretrial sentencing determinations, but it lacked governance about proper oversight, calculating risk levels, or methods to protect against biased outcomes.<sup>169</sup>

A super-regulation may resolve some of these concerns by governing AI technology from its initial inception in the IRB review process (e.g., ensuring the training data is unbiased), through FDA clearance (e.g., testing to ensure quality and nondiscriminatory effects), all the way up to implementation and use by healthcare entities (e.g., appropriate review by providers and compliance with Section 1557). Nicholas Terry proposes three characteristics of a regulatory matrix for AI: Unitary, holistic, and universal.<sup>170</sup> These characteristics applied to the nondiscrimination principles addressed here can create an entirely new regulatory regime for healthcare AI:

- Unitary. Rather than separating AI device and practice of medicine using AI, a regulatory framework can address both in tandem;
- Holistic. A regulatory framework for AI can address the nondiscrimination concerns discussed here, as well as other criticisms (e.g., quality and safety, transparency, data protection, and cost effectiveness);
- Universal. A universal approach will ensure that clinical and administrative uses of AI are not treated differently, and that different healthcare entities (e.g., hospitals and insurers) are not treated differently. Relatedly, the super-regulation would apply to all healthcare entities regardless of whether they receive federal funds. Given the growing use of AI and potential for data-sharing, a universal regulation could also protect nondiscrimination if the data contained in AI leaves the healthcare domain.

---

<sup>167</sup> See *id.* at 173; see also Sandra Wachter et al., *Why a Right to Explanation of Automated Decision-Making Does Not Exist In the General Data Protection Regulation*, 7 INT’L. DATA PRIVACY L. 76 (2017) (Europe).

<sup>168</sup> Tsukayama & Williams, *supra* note 161.

<sup>169</sup> *Id.*

<sup>170</sup> Terry, *supra* note 9, at 163–64.

## VII. CONCLUSION

*“Against this historical backdrop [of racial imbalance], it is imperative that pretrial risk assessment instruments, if used at all, be designed to help meet the goal of reducing racial disparities . . . . If a tool cannot help achieve that goal, then it is not a tool that the justice system needs.”*<sup>171</sup>

The quote above is from a statement released by over 100 advocacy organizations in response to California Senate Bill 10, a legislation that essentially approved the use of AI in the criminal justice context.<sup>172</sup> As emphasized throughout this Article, AI-informed decision-making is emerging in nearly every industry. From determinations made by the government (e.g., public benefits) to those made by private actors (e.g., employers), researchers and advocates are concerned about the discriminatory effects inherent in AI. AI used in the healthcare context, however, presents a unique challenge: these biases can permeate the intimate nature of the examination room or operating room.

This Article discussed the effects of AI bias from a legal and compliance standpoint. The existing nondiscrimination responsibilities on healthcare entities through Section 1557 of the ACA are likely to extend to healthcare AI. Similarly, licensing requirements and the threat of malpractice liability will continue to hold physicians to a standard of care when using AI, one that does not condone discrimination. However, those responsibilities do not govern the design and development of AI—where bias is most likely to be imbedded. As discussed in this Article, the FDA’s enforcement authority is currently too limited to adequately oversee all types of healthcare AI and is unlikely to address discrimination.

The promise of AI is attractive: Faster, less costly, and more accurate decision-making. Advocates of healthcare AI technology argue that it has the potential to reduce errors resulting from human variation by physicians.<sup>173</sup> In fact, these advocates argue that AI has the potential to *eliminate*, rather than introduce or perpetuate, biases in healthcare because it can be programmed to not be influenced by external information about patients or their finances.<sup>174</sup> AI even presents a novel opportunity to remedy many health disparities. For example, AI

---

<sup>171</sup> STATEMENT ON RISK ASSESSMENT TOOLS, *supra* note 66, at 3.

<sup>172</sup> See Matthew Borges, *California rejects proposition to end cash bail*, JURIST (Nov. 5, 2020), <https://www.jurist.org/news/2020/11/california-rejects-proposition-to-end-cash-bail/> (noting that the bill was never implemented, it was stayed pending a statewide voter referendum and California voters rejected it on November 3, 2020)

<sup>173</sup> See Jennifer Bresnick, *Unleashing the Value of Health Data in the Era of Artificial Intelligence*, HEALTH ANALYTICS (Apr. 15, 2019), <https://healthitanalytics.com/features/unleashing-the-value-of-health-data-in-the-era-of-artificial-intelligence> (noting that AI used in radiology can detect subtleties in tissue structure that is often missed by the human eye).

<sup>174</sup> See *id.*

could be used as a tool in psychiatry to diagnose or treat individuals with severe mental illness.<sup>175</sup>

But the “promise” of AI is misleading. Without a comprehensive (legislative, regulatory, or industry standard) framework that addresses biases in AI, patients that have historically not benefited from the healthcare industry will continue to face discrimination—engrained systemic biases, will only become solidified, automated ones. Patients belonging to suspect classes or low-socio-economic communities have historically been excluded from the benefits of the American healthcare system,<sup>176</sup> and patients have little reason to trust to automation.<sup>177</sup> As AI explodes as a clinical and administrative tool, nondiscrimination principles should be at the center of regulators’ and the industries’ plans.

---

<sup>175</sup> Jamie Ducharme, *Artificial Intelligence Could Help Solve America’s Impending Mental Health Crisis*, TIME (Nov. 20, 2019), <https://time.com/5727535/artificial-intelligence-psychiatry/> (noting that some AI can detect abnormalities in speech patterns that humans can never perceive, patterns that indicate schizophrenia).

<sup>176</sup> See Terry, *supra* note 9, at 170. Terry argues the “moral imperative of societal good” is more of a political question than an ethical one. He notes that the United States has not “embraced the healthcare solidarity that typically underpins such discussions about societal good.”

<sup>177</sup> Michael Mittelman, Sarah Markham & Mark Taylor, *Patient Commentary: Stop Hyping Artificial Intelligence--Patients Will Always Need Human Doctors*, 363 BMJ k4669 (2018). (“Patients haven’t always benefited from the promises of technology . . . Technology companies have given patients few reasons to trust them with all their medical data.”).