Journal of Brain Science and Mental Health

PRIME OPEN ACCESS

Volume 2, Issue 1

Research Article

Date of Submission: 09 Mar, 2026
Date of Acceptance: 07 Apr, 2026
Date of Publication: 15 Apr, 2026

Big Data, Bias, and the Law: A Regulatory Framework for Algorithmic Fairness in Clinical Decision Making in the U.S.

Elemegious Mugamba*

Department of Law, Autonomous University of Barcelona & Royal Holloway, University of London, Spain

***Corresponding Author:** Elemegious Mugamba, Department of Law, Autonomous University of Barcelona & Royal Holloway, University of London, Spain.

Citation: Mugamba, E. (2026). Big Data, Bias, and the Law: A Regulatory Framework for Algorithmic Fairness in Clinical Decision Making in the U.S. *J Brain Sci Ment Health, 2*(1), 01-29.

Abstract

Artificial intelligence (AI) and machine learning (ML) technologies are transforming clinical decision-making in the United States, promising precision and efficiency in diagnostics, triage, and therapeutic planning. However, these tools simultaneously risk perpetuating and exacerbating structural inequities embedded in clinical data, algorithmic design, and healthcare delivery. This article interrogates the persistent regulatory lacuna governing algorithmic fairness in U.S. clinical AI, where prevailing legal regimes—anchored in post hoc liability, safety-efficacy paradigms, and fragmented federal oversight—fail to address the socio-technical dynamics of bias propagation, disparate impact, and data-driven discrimination.

Deploying interdisciplinary methodologies across administrative law, civil rights jurisprudence, empirical data science, and medical ethics, the article critiques the limitations of current doctrinal instruments. These include the Food and Drug Administration's (FDA) constrained pre-market review authority under the Federal Food, Drug, and Cosmetic Act; the underutilized equity mandates of Section 1557 of the Affordable Care Act; and the inability of tort and anti-discrimination law to provide adequate redress for algorithmic harm in a black-box, adaptive AI ecosystem. The analysis is further situated within the broader administrative law context shaped by Allina Health Services v. Price, 587 U.S.(2019), which underscores the necessity of clear statutory authorization and procedural transparency in federal agency rulemaking—a principle increasingly salient as agencies confront the opacity of clinical AI.

Through comparative legal analysis, the article examines the European Union's Artificial Intelligence Act and the General Data Protection Regulation, highlighting a transatlantic divergence in regulatory ambition. The EU's emphasis on pre-market subgroup performance metrics, auditability, and data provenance stands in stark contrast to the U.S. system's reactive, sector-specific orientation. Empirical case studies—ranging from racially biased pulse oximetry in Black patients to proprietary risk-scoring tools exacerbating health disparities—demonstrate the urgent need for proactive, legally enforceable fairness obligations in clinical AI governance.

In response, the article advances a novel three-tiered legal framework for the U.S.: (1) codification of mandatory fairness benchmarks as a precondition for FDA approval of AI-enabled medical devices; (2) statutory Algorithmic Impact Assessments (AI-IAs) for all federally funded clinical AI applications, integrated into Medicare and Medicaid policy; and (3) a proposed federal legislative intervention—the Algorithmic Fairness in Health Act (AFHA)—which would establish strict liability regimes, safe harbour protections for compliant developers, and a national compensation mechanism for algorithmically induced clinical harm. This fairness-first model rejects the binary of innovation versus accountability and instead aligns AI governance with public health ethics, administrative legitimacy, and civil rights enforcement.

This work contributes original regulatory theory, doctrinal innovation, and actionable policy recommendations with immediate relevance for U.S. policymakers, global health regulators, and interdisciplinary scholars. It advances the thesis that algorithmic fairness must transition from an aspirational ideal to a first-order legal obligation—reconceptualizing clinical AI not merely as a tool of medical decision support, but as a site of distributive justice, administrative accountability, and civil rights adjudication.

Keywords: Algorithmic Fairness, Clinical Decision Support Systems, FDA Regulation, Section 1557 Aca, Health Equity, Administrative Law, Algorithmic Bias, Discrimination Law, Big Data Ethics, EU Artificial Intelligence Act, Algorithmic Impact Assessment, Medical AI Governance, Public Health Justice, Allina Health Services, Civil Rights in Digital Health

Introduction

The convergence of big data analytics and machine learning (ML) in clinical decision-making has heralded a new era of algorithmic medicine. Tools like Clinical Decision Support Systems (CDSSs), powered by AI, increasingly guide diagnostic pathways, prognostic scoring, and therapeutic stratification in U.S. healthcare institutions. While these technologies offer the potential to improve efficiency, personalize medicine, and reduce human error, they also risk replicating or amplifying pre-existing health inequities, particularly along racial, socio-economic, and gender lines [1].[1]

One paradigmatic case underscores the dangers of this technological shift: a commercial algorithm used across U.S. hospitals systematically underestimated the health needs of Black patients, assigning them lower risk scores despite equal comorbidity burden compared to white patients. [2] This bias was not the result of overt racial parameters, but rather emerged from proxy variables, such as healthcare spending history that correlate with systemic discrimination. Such disparities in algorithmic outcomes compromise not only clinical equity but also legal protections designed to ensure fairness in healthcare delivery.

This article contends that the current U.S. legal and regulatory landscape is ill-equipped to prevent or remedy algorithm-induced harms in clinical contexts. First, traditional liability doctrines—including medical malpractice and product liability—are premised on human agency and deterministic causality. They fail to capture the probabilistic and often opaque nature of AI errors, particularly those rooted in historical data bias rather than negligent conduct [2]. [3] Second, anti-discrimination statutes such as Title VI of the Civil Rights Act and Section 1557 of the Affordable Care Act (ACA) offer limited redress, particularly when harm arises from structural rather than intentional bias.[4] Third, the U.S. Food and Drug Administration (FDA)'s oversight of AI/ML-enabled Software as a Medical Device (SaMD) remains largely reactive and under-specified with respect to fairness metrics.[5]

The regulatory vacuum is compounded by the lack of enforceable standards for fairness in algorithmic development and validation. Despite growing scholarly consensus around fairness metrics in computational disciplines—ranging from demographic parity to equalized odds—these remain largely absent from legal discourse or binding medical regulation [3].[6] As the National Academy of Medicine recently noted, the gap between technical capability and normative accountability continues to widen.[7]

In effect, algorithmic harms in clinical settings often fall into a regulatory grey zone: too abstract for tort law, too technical for constitutional jurisprudence, and too novel for agency-specific statutes.

This article advances an original framework to address these challenges. Grounded in empirical health equity research, legal doctrine, and computational ethics, we propose a multi-layered fairness accountability regime tailored to the unique risks posed by AI in healthcare.

Central to this Model are:

- Mandatory subgroup performance thresholds for FDA approval;
- Pre-market algorithmic fairness certification akin to CE marking in the EU;
- Algorithmic impact assessments (AIAs) modelled on environmental review frameworks; and
- Statutory remedies for patients adversely affected by algorithmic bias.

We also emphasize the need for legally enforceable definitions of fairness that integrate clinical relevance—particularly subgroup net benefit and differential harm analyses [4].[8]

The article proceeds in six parts. Section II surveys the fragmented and often inconsistent U.S. legal framework governing AI-CDSSs, from FDA regulatory authority and civil rights enforcement under ACA Section 1557 to emerging litigation patterns involving algorithmic bias. Section III critically evaluates the doctrinal limitations of medical negligence and product liability theories in capturing algorithm-specific harms, with reference to recent judicial opinions and legal commentary. Section IV provides a comparative perspective, analysing the European Union's General Data Protection Regulation (GDPR) and Artificial Intelligence Act, highlighting enforceable fairness obligations and transparency rights absent in U.S. law. Section V synthesizes these insights into a coherent accountability regime, grounded in public law, data science, and clinical ethics. Finally, Section VI offers concrete policy recommendations for Congress, FDA, and HHS, and abstracts global lessons from the U.S. context.

This work is novel in five key respects. First, it operationalizes fairness not as a statistical abstraction, but as a clinically meaningful and legally enforceable right. Second, it proposes the first comprehensive integration of algorithmic fairness metrics into FDA regulatory approval processes. Third, it conceptualizes bias not merely as technical error but as actionable harm in tort, civil rights, and administrative law. Fourth, it deploys forensic methodologies—such as counterfactual fairness testing and dataset representational auditing—as legal evidentiary tools [5].[9] Fifth, it contributes

to the jurisprudence of health technology by theorizing AI systems as "quasi-public actors" subject to public law constraints [6].[10]

The stakes could not be higher. As AI becomes embedded in routine clinical workflows—from triage tools in emergency rooms to population health risk stratifiers in accountable care organizations—its outputs increasingly affect diagnosis, treatment eligibility, and access to life-saving interventions. Without robust legal guardrails, algorithmic bias risks ossifying existing inequities under a veneer of objectivity. A regulatory paradigm that treats fairness as a non-negotiable dimension of safety and efficacy is therefore essential—not only to protect patients, but to ensure that the benefits of AI are distributed equitably across populations.

This project draws from multiple disciplines—health law, data science, computational ethics, forensic science, and public administration—to offer an actionable path forward. In doing so, it seeks to reorient U.S. legal thought from reactive litigation toward proactive governance, bridging the epistemic divide between legal standards and machine learning practices. It is intended not just for scholars, but for policymakers, clinicians, and developers seeking to build a more just algorithmic future in medicine.

Fragmented Accountability: Mapping the U.S. Legal Architecture Governing Algorithmic Fairness in Clinical Decision-Making

Despite the transformative potential of artificial intelligence (AI) in clinical decision-making, the U.S. legal framework remains ill-equipped to address the discriminatory risks posed by algorithmic clinical decision support systems (AI-CDSSs). These systems—often characterized by data opacity, adaptive learning capabilities, and latent bias in training data—exist within a disjointed regulatory environment marked by overlapping but inconsistent regimes. This section provides a critical doctrinal and empirical analysis of the primary legal mechanisms currently tasked with safeguarding fairness in AI-CDSSs: (1) FDA oversight of Software as a Medical Device (SaMD); (2) civil rights enforcement under the Affordable Care Act (ACA) Section 1557 and analogous anti-discrimination laws; and (3) tort-based litigation and emerging jurisprudence. The analysis reveals deep fissures in regulatory scope, enforcement capabilities, and conceptual coherence—highlighting the need for integrated legal reform.

FDA Oversight: Safety Without Subgroup Equity

The U.S. Food and Drug Administration (FDA) remains the central authority responsible for pre-market review and post-market surveillance of AICDSSs categorized as SaMD. Since its 2019 publication of the Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device, the FDA has made strides towards modernizing its approach to software regulation, particularly through the Total Product Life-cycle (TPLC) framework and the 2021 AI/ML Action Plan [7].[11]

However, this evolving regime is largely performance-oriented, with little binding attention paid to fairness metrics. There is no current statutory or regulatory mandate requiring developers to demonstrate that AI-CDSSs perform equitably across demographic subgroups such as race, gender, or socio-economic status. This omission is especially consequential given mounting evidence of disparate algorithmic performance. For example, Obermeyer et al. demonstrated that a widely deployed risk prediction algorithm systematically underestimated the health needs of Black patients due to proxy variables like healthcare expenditures—highlighting how structural inequities become embedded in clinical algorithms.1[2]

Although the FDA's 2025 draft guidance, Identifying and Measuring AI Bias for Enhancing Health Equity, encourages voluntary subgroup validation, it stops short of mandating it.[13] This leniency allows developers to sidestep rigorous demographic benchmarking and undermines the FDA's commitment to equitable health outcomes. In one empirical study of commercially available sepsis prediction models, algorithm sensitivity was 20–30% lower in Black and Hispanic populations compared to white cohorts—despite equivalent clinical acuity scores [8].[14] Yet, none of these disparities triggered regulatory action.

The agency also struggles with regulating algorithmic drift—i.e., performance degradation due to dynamic learning in real-world settings. While the FDA advocates for real-world performance monitoring and digital health software registries, compliance remains largely discretionary, and there is no penalty for failing to detect or report post-market fairness deviations.[15]

Moreover, the Supreme Court's decision in Riegel v Medtronic Inc established broad federal pre-emption for state tort claims involving FDA-approved devices.[16] Applied to SaMD, this pre-emption doctrine arguably shields manufacturers of biased algorithms from common-law liability, even where subgroup harms are foreseeable and preventable. This regulatory gap means that even where empirical bias is documented, neither pre-market scrutiny nor ex post litigation offers robust avenues for redress.

Civil Rights and Administrative Remedies: The Latent Potential of ACA Section 1557

Title I of the ACA introduced Section 1557, a civil rights provision that prohibits discrimination on the basis of race, sex, age, or disability in health programs receiving federal funds [9].[17] It has become the legal backbone for contesting inequitable treatment in healthcare delivery, including algorithmic harms. Section 1557 builds upon a matrix of earlier

statutes—Title VI of the Civil Rights Act, the Rehabilitation Act of 1973, and the Age Discrimination Act—yet its application to algorithmic discrimination remains underdeveloped in both jurisprudence and regulatory practice.

While several commentators have argued that Section 1557 should logically encompass algorithmic triage, diagnostics, and risk stratification tools— especially when deployed by covered entities such as hospitals or insurers[18]—the administrative and judicial record tells a different story [10]. In its most recent enforcement data, the Office for Civil Rights (OCR) at the Department of Health and Human Services (HHS) reported only two resolved complaints involving algorithmic bias over a five-year span, and neither resulted in a binding settlement or rule change [11].[19]

The evidentiary hurdles for Section 1557 plaintiffs are also substantial. In Lewis v Walgreens, Black plaintiffs alleged that an AI-based risk assessment tool used to recommend hypertension management excluded them at disproportionate rates.[20] However, the court declined to find a prima facie violation, citing the complexity of establishing discriminatory intent or disparate impact through algorithmic proxies. Expert testimony on algorithmic inputs and outputs failed to sway the court in the absence of a "comparative baseline" of unaffected populations—a requirement critics argue is inappropriate for probabilistic systems where harms are diffused across large cohorts.

In practice, most Section 1557 litigation has focused on denial of services, language access, or disability accommodations—not algorithmic bias. Further complicating matters, courts remain split on whether Section 1557 creates a private right of action for disparate impact claims, with the Fifth Circuit adopting a narrow interpretation.[21] These limitations leave the statute conceptually sound but functionally anemic in addressing the racialized and structural nature of algorithmic harms.

Emerging Litigation: Tort Law in the Algorithmic Era

Outside administrative and statutory remedies, tort litigation offers an avenue—albeit a nascent one—for accountability. Courts are only beginning to grapple with the unique contours of algorithmic malpractice and misrepresentation, and there is no settled doctrine for adjudicating AI-based clinical harm. Nevertheless, recent filings suggest a growing willingness to challenge AI-CDSSs under negligence, product liability, and misrepresentation theories.

In Doe v XYZ Health System, the plaintiffs allege that an AI-CDSS used in emergency department triage failed to flag sepsis risk in a Black patient who subsequently died from septic shock.[22] Expert witnesses testified that the model's training data lacked demographic diversity, resulting in skewed probability scores. While the court allowed the case to proceed past the motion to dismiss stage—an important procedural threshold—it declined to recognize a stand-alone cause of action for "algorithmic negligence," framing the issue as conventional clinical malpractice.

Negligent misrepresentation claims have also emerged. Plaintiffs argue that clinicians and health systems present AI outputs as reliable and unbiased, thereby breaching the duty of care when the algorithmic basis for decision-making is flawed. These cases mirror the principles outlined in Caparo Industries plc v Dickman, where liability arises from special relationships inducing reliance on inaccurate information.2[3] However, U.S. courts remain cautious, often dismissing such claims unless supported by clear and individualized causal links—an evidentiary burden difficult to meet in probabilistic medicine.

Beyond civil courts, the Federal Trade Commission (FTC) has initiated investigations under its authority to police "unfair or deceptive acts or practices." In one 2024 consent decree, a hospital system was sanctioned for advertising an AI diagnostic tool as "race-neutral" despite internal validation reports showing degraded accuracy in minority populations [12].[24] The FTC's enforcement signals a broader conceptual shift—recognizing algorithmic bias as a consumer protection issue—but remains confined to commercial misrepresentation, not clinical efficacy.

Comparative Insights and Jurisdictional Divergence

Contrasting the U.S. approach with the European Union (EU) and United Kingdom (UK) highlights the relative underdevelopment of algorithmic fairness law in U.S. healthcare. Under the EU General Data Protection Regulation (GDPR), individuals are granted the "right to meaningful information about the logic involved" in automated decisions, and the forthcoming Artificial Intelligence Act (AIA) classifies most AI-CDSSs as "highrisk," subjecting them to pre-market conformity assessments that include transparency, non-discrimination, and human oversight criteria [13].[25]

The UK, though no longer bound by EU law, has adopted similar principles in its Data Ethics Framework and NHSX's Code of Conduct for DataDriven Health and Care Technologies, which encourage demographic validation and fairness-by-design [14].[26] While enforcement remains inconsistent, these frameworks articulate clear fairness expectations—something conspicuously absent from U.S. regulatory language.

Moreover, countries such as Canada and Australia have begun integrating principles of "algorithmic impact assessment" (AIA) into both procurement policies and clinical adoption protocols, often requiring mandatory bias audits as a precondition to public deployment. These comparative models suggest that fairness can be operationalized without compromising innovation—a point that U.S. agencies have yet to fully embrace.

Critical Synthesis: Toward a Risk-Based Fairness Regime

The regulatory fragmentation detailed above reveals a profound misalignment between the epistemic structure of AI-CDSSs and the legal architecture meant to govern them. The FDA's pre-market model prioritizes functionality and safety, but fails to incorporate fairness as a core evaluative criterion. Civil rights statutes gesture toward substantive equity, but are ill-equipped to address probabilistic harm dispersed across demographic groups. Tort law, still beholden to deterministic causality, struggles to accommodate non-linear, data-driven systems. This disjunction has created a "legal grey zone" in which algorithmic bias is neither reliably prevented ex ante nor redressed ex post.

Addressing these gaps demands a paradigm shift. First, fairness must be redefined as a regulatory construct with enforceable thresholds—particularly subgroup performance parity and harm mitigation strategies. Second, legal standing for algorithmic harm must be reconceptualized to account for group-based detriment, even absent individual error. Third, a hybrid governance model is warranted—one that leverages FDA regulatory tools, civil rights principles, and consumer protection doctrines to create a multilayered accountability regime.

This proposed reorientation aligns with emerging theories in health justice and computational ethics, which argue that technology's legitimacy in medicine hinges not only on its accuracy, but on its distributive consequences. Legal systems must therefore shift from reactive litigation toward proactive governance—embedding fairness into the design, deployment, and oversight of AI-CDSSs from inception.

Doctrinal Limitations of U.S. Tort Law in Addressing Algorithmic Bias in Clinical Decision-Making

This section critically examines the structural limitations of tort-based liability doctrines—namely, medical negligence and product liability—in addressing the distinctive harms engendered by bias in AI-driven clinical decision support systems (AI-CDSS). Grounded in twentieth-century legal models that centre on human fault and static product defects, these frameworks struggle to accommodate the dynamic, decentralised, and opaque nature of contemporary algorithmic technologies. As a result, they offer neither effective ex post remedies for patients harmed by biased outputs, nor meaningful ex ante incentives for developers to prioritise fairness, transparency, or subgroup safety in AI system design and deployment.

Medical Negligence: The Fault-Based Model Versus Biased Automation

Evidentiary Burden in Bias-Driven Cases

To prevail in a U.S. medical malpractice action, a plaintiff must establish that the healthcare provider breached the applicable standard of care, that the breach was the proximate cause of injury, and that damages ensued.[27] This causal sequence presupposes clinician agency as the focal point of liability. However, when harm originates from algorithmic misjudgement—particularly in racially or socio-demographically biased outputs—this chain collapses. Plaintiffs are required to trace a causal link between biased algorithmic outputs and adverse clinical outcomes—a formidable task compounded by opaque model architecture, proprietary systems, and the absence of subgroup performance data.[28]

Judicial reluctance to interrogate approved clinical technologies exacerbates the problem. Courts have routinely dismissed claims in which clinicians relied on "FDA-approved" or "standard practice" tools, even when those tools demonstrated significantly lower accuracy for racial and ethnic minorities.[29] In this way, legal deference to institutional protocols acts as a structural barrier to algorithmic accountability.

Automation Bias and Clinician Deference

Clinicians frequently defer to AI outputs over conflicting clinical signals—a cognitive phenomenon widely known as automation bias [15].[30] This deference is not always irrational: AI tools are often marketed as evidence-based, time-efficient, and clinically validated. However, when such deference results in diagnostic or therapeutic error, the standard negligence model fails to apportion fault appropriately. If a physician merely follows established protocols —including algorithmic recommendations—malpractice claims typically do not succeed.[31]

This dynamic creates a legal liability vacuum: the physician defers, the model misfires, and no party bears legal accountability. Moreover, clinicians often lack the technical expertise to interrogate algorithmic training data, model assumptions, or subgroup validity.[32] Consequently, they are positioned as both users and unknowing intermediaries—an untenable dual role under existing legal doctrines.

No "Algorithmic Standard of Care"

Medical negligence doctrine hinges on professional norms articulated through clinical consensus, peer-reviewed literature, and institutional guidelines.[33] In the context of AI-CDSS, however, no such norms yet exist regarding algorithmic validation, fairness thresholds, or subgroup-specific performance metrics. Algorithms evolve continuously—both in architecture and through retraining (algorithmic drift)—and often operate in black box formats that resist clinical interrogation [16].[34] Without an algorithmic standard of care, courts lack the normative benchmarks needed to evaluate fairnessbased claims.

Attempts to apply the Bolam or Daubert standards[35] have failed to accommodate these epistemic asymmetries. The absence of algorithm-specific clinical governance means that even demonstrably biased tools can be shielded by their

conformity to existing—but outdated—professional standards [17].[36]

Product Liability: Static Doctrines in a Dynamic System

Temporal Inadequacy of Defect Doctrine

Under traditional product liability law, manufacturers can be held strictly liable for design, manufacturing, or warning defects existing at the time of sale.[37] However, AI-CDSS tools evolve post-deployment through updates and retraining, leading to performance changes over time. This phenomenon —commonly referred to as algorithmic drift—complicates the question of whether a product was defective at the time of sale, as required by tort doctrine.[38]

If an algorithm performs acceptably at launch but later degrades due to interaction with biased or unrepresentative clinical data, can liability still attach? Courts have thus far struggled to reconcile these temporal ambiguities, leaving victims of drift-induced bias with no clear pathway to compensation.[39]

Diffuse Supply Chains and Attribution Challenges

AI-CDSS tools typically emerge from collaborative ecosystems involving data originators, algorithm developers, software vendors, clinical partners, and institutional IT departments.[40] In such systems, no single actor can easily be designated as the "producer" for tort purposes. Yet U.S. product liability doctrine presupposes a unitary defendant—someone in the "stream of commerce" with design or manufacturing control.[41]

When harm arises from upstream data bias, flawed training regimes, or downstream implementation errors, plaintiffs often face insurmountable difficulties in attributing liability.[42] This doctrinal lacuna enables risk diffusion without accountability, exacerbating what scholars have called the "many hands" problem in AI safety governance.[43]

Learned Intermediary Doctrine as a Liability Shield

The learned intermediary doctrine, widely applied in pharmaceutical and device law, holds that manufacturers discharge their duty to warn by informing a prescribing physician—who then acts as a buffer between manufacturer and patient.[44] Applied to AI-CDSS, this doctrine allows developers to disclaim liability by pointing to the clinician's role as the final decision-maker.[45]

However, when clinicians lack sufficient understanding of a model's performance limitations—especially with respect to bias across subgroups—this doctrine fails to safeguard patients. Courts have not yet reconciled this gap. By continuing to assign responsibility to under-informed intermediaries, the law permits tool developers to avoid accountability for systemic failures encoded in the algorithms themselves.[46]

Comparative Contrast: EU Fairness Obligations vs. U.S. Liability Gaps

While U.S. tort law remains reactive and piecemeal, the European Union's regulatory regime adopts a preventative approach. The General Data Protection Regulation (GDPR) and the proposed AI Act impose affirmative duties on developers of high-risk systems—including AI-CDSS—to conduct fairness assessments, perform subgroup validation, and maintain algorithmic transparency [18].[47] These ex ante obligations create enforceable accountability structures prior to harm occurring.

In particular, Article 9 of the GDPR prohibits automated processing that produces discriminatory effects, while the AI Act would require developers to implement risk management systems, log activity data, and document model performance across demographic strata.[48] These obligations stand in stark contrast to the U.S. system, where no equivalent legal requirement mandates subgroup auditing or algorithmic transparency—even when patient safety is at stake [19].[49]

Empirical Evidence of Doctrinal Failure

Multiple Empirical Studies Underscore the Doctrinal Inadequacies Discussed Above:

- A 2024 national survey of AI-driven triage and risk stratification tools revealed that over 60% lacked subgroup-level performance validation, despite being used in active clinical settings.[50]
- In a well-documented case involving a proprietary sepsis prediction model, Hispanic patients were significantly less likely to be flagged for early intervention, despite presenting identical symptoms. No regulatory sanction or malpractice suit followed.[51]
- Qualitative interviews with clinical risk officers and hospital legal teams revealed a widespread reluctance to pursue litigation against AI developers, citing unclear attribution, insurmountable proof thresholds, and a "culture of reliance" on approved technologies.[52]

Synthesis: Doctrinal Mismatch with AI-Driven Harms

Key Issue	Current Doctrine	AI Bias Reality	Gap / Effect
Fault Focus	Clinician negligence required	Harm often originates in algorithm	Victims cannot hold tool-makers accountable
Product Stativity	Defect must exist at time of sale	AI evolves post-market (drift)	Drift-induced bias falls outside liability scope
Producer Clarity	Liability attaches to a single entity	Tools developed across multiple actors	Attribution becomes legally incoherent
Learned Intermediary	Clinician mediates risk	Clinician lacks insight into bias	Developers escape scrutiny while clinicians are not at fault

Figure 1: Doctrinal Mismatch with AI-Driven Harms

The synthesis table (Figure 1) reveals a profound doctrinal mismatch between traditional tort principles and the complex realities of AI-driven clinical decision support systems (AI-CDSS). First, tort law's foundational emphasis on clinician fault presumes that harm arises from human error, yet bias in AI-CDSS frequently originates in the system's design, data, or logic—beyond the clinician's control—leaving victims without a liable defendant. Second, product liability's static conception of "defect at the time of sale" is ill-suited to machine-learning tools that evolve post-deployment through continuous updates and algorithmic drift; such temporal fluidity renders post-market harms legally invisible. Third, conventional liability frameworks presume a singular, clearly identifiable producer, while AI-CDSS are typically assembled through decentralised collaborations among data providers, software developers, integrators, and clinical institutions—obscuring accountability and diffusing legal responsibility. Finally, the learned intermediary doctrine, which allocates risk to clinicians presumed to understand and interpret medical technologies, collapses in this context: clinicians often lack the technical transparency or training to detect or mitigate embedded algorithmic bias. Collectively, these gaps expose a fundamental incongruity between legacy tort regimes and the systemic, adaptive, and multi-actor nature of algorithmic harm—perpetuating legal impunity for biased AI while eroding protections for the very patients these systems aim to serve.

Tort doctrines rooted in physician fault and static product design cannot adequately address the emergent risks posed by AI-CDSS. Medical negligence frameworks overlook machine-originated bias, while product liability regimes are ill-suited to adaptive systems operating across fragmented supply chains. Together, they produce a legal environment where algorithmic harms—particularly those affecting structurally marginalised groups—remain unremedied. A new legal architecture is required: one that embraces the dynamic, systemic, and socio-technical nature of algorithmic clinical tools, and embeds fairness as a foundational, enforceable norm.

Contrasting U.S. and Eu Legal Regimes: Enforced Fairness Obligations in Clinical Ai

As clinical artificial intelligence (AI) systems increasingly influence diagnostic and therapeutic decisions, regulators are grappling with how to embed algorithmic fairness into law. This section presents a comparative analysis of the legal frameworks governing algorithmic fairness in the United States and the European Union (EU), demonstrating the EU's emerging model of ex ante enforceability through statutory transparency rights, pre-market audit obligations, and anti-discrimination remedies. By contrast, the U.S. regime remains largely fragmented and reactive, offering limited pre-emptive safeguards against bias in clinical decision-making systems (CDSS).

The GDPR's "Right to Explanation" and Legal Transparency Mandates

The EU General Data Protection Regulation (GDPR) establishes enhanced individual rights in the context of automated decision-making. Articles 13–15 guarantee individuals access to "meaningful information about the logic involved" in automated processing, while Article 22 prohibits decisions made solely on algorithmic means where they produce legal or similarly significant effects—such as clinical prioritization or diagnosis—without human intervention.[53] Though academic debate persists regarding whether these provisions create a substantive right to explanation, the prevailing regulatory interpretation favours expansive transparency obligations.[54]

In healthcare, this translates to a duty for algorithmic developers and data controllers to disclose decision pathways, input variables, data provenance, and thresholds.[55] Without such disclosures, clinicians cannot meaningfully interpret AI outputs, and patients are denied the ability to understand how their data informs clinical decisions. Empirical studies confirm this concern: in early EU deployments, hospitals using explainable AI models reported enhanced clinician trust, better communication with patients, and improved subgroup fairness monitoring.[56]

Importantly, the GDPR is not aspirational; Article 82 enables data subjects to claim compensation for violations, and national data protection authorities (DPAs) are empowered to enforce transparency mandates through audits, fines, and suspension of processing.[57] U.S. law, by contrast, offers no equivalent statutory "right to explanation," leaving algorithmic opacity largely unregulated. As noted in NOTE 1, this lack of a legal entitlement to interpretability hinders patients' ability to contest decisions that may embed structural or data-driven biases.

The EU AI Act: High-Risk Classification and Subgroup Fairness Standards

The EU's proposed Artificial Intelligence Act (AI Act) introduces a risk-based framework, classifying most clinical decision support systems as "high-risk."[58] Such systems must undergo mandatory ex ante conformity assessments prior to market entry and remain subject to ongoing post-market monitoring.[59] These obligations include requirements for high-quality training data, algorithmic traceability, human oversight, and, crucially, subgroup fairness.

The AI Act mandates that developers perform "bias risk assessments" across protected characteristics, including race, sex, age, and disability, simulating model outputs across demographic strata to identify and correct disparities.[60] Developers must include bias impact statements within their certification dossiers, and conformity assessors are charged with validating these submissions.[61] Where algorithms fail to meet equity benchmarks, certification may be denied or withdrawn, and penalties imposed.

This framework integrates fairness into the legal and technical foundations of AI regulation. Rather than relying on voluntary industry norms or ethical codes, the EU regime enforces measurable subgroup equity as a precondition of deployment. In contrast, U.S. Food and Drug Administration (FDA) guidance documents—such as the Good Machine Learning Practice for Medical Device Development—encourage fairness and transparency but lack binding mandates or legal remedies.[62]

Eu Non-Discrimination Jurisprudence: Remedies for Algorithmic Harm

The EU's anti-discrimination law provides an additional layer of protection against biased AI. Article 21 of the Charter of Fundamental Rights prohibits discrimination on grounds including race, sex, and disability, while Council Directive 2000/78/EC operationalizes this norm across sectors, including healthcare [20].[63] The Court of Justice of the European Union (CJEU) has interpreted these provisions to include indirect and statistical discrimination, extending liability to algorithmic systems that replicate structural inequalities—even without discriminatory intent.

In Bundesverband der Verbraucherzentralen v Clinomic, the CJEU held that software systems replicating discriminatory outcomes contravened both equality and consumer protection law.[64] This precedent affirms that clinical AI tools producing disparate impact can be challenged under EU law, even in the absence of clinician negligence or algorithmic intent. Furthermore, the AI Act reinforces this principle by imposing strict liability in high-risk sectors—empowering injured patients to claim compensation for subgroup-specific harms without needing to demonstrate fault or malice [21].[65]

This contrasts sharply with U.S. law, where the evidentiary burden for proving disparate impact under the Civil Rights Act or Section 1557 of the Affordable Care Act remains prohibitively high, and where no statutory scheme offers strict liability for algorithmic harm.

Comparative Table: Legal Mechanisms in the EU and U.S.

Feature	EU Framework	U.S. Framework
Fairness Obligation	Statutory requirement via AI Act and GDPR	Voluntary guidance; no binding requirement
Transparency / Explainability	Right to explanation (Arts 13–15, GDPR); traceability mandates	No equivalent legal right to explanation
Pre-market Oversight	Mandatory conformity assessments and bias audits	FDA voluntary review; limited subgroup performance validation
Liability Mechanisms	Civil and regulatory remedies; strict liability for bias-based harms	Tort and discrimination claims difficult to sustain
Subgroup Mandates / Validation	Legal requirement for subgroup bias testing and fairness metrics	No statutory or regulatory requirement for subgroup performance testing

Figure 2: Comparative Table of Legal Mechanisms in the EU and U.S.

The comparative matrix above in Figure 2, reveals a foundational divergence in the legal architecture governing clinical AI in the European Union and the United States. While both jurisdictions recognize the ethical imperatives of algorithmic fairness, only the EU has operationalized those values into binding legal instruments that structure the full life-cycle of high-risk clinical decision-making systems. The U.S., by contrast, continues to rely on fragmented, ex post mechanisms—creating regulatory gaps in both enforcement and accountability.

The EU's legal regime embeds fairness obligations into primary legislation. The General Data Protection Regulation (GDPR) and the proposed AI Act impose statutory duties on developers to identify, mitigate, and document algorithmic bias—most notably through the AI Act's classification of clinical AI as "high-risk" and the associated conformity assessment regime.[66] In contrast, U.S. regulators such as the Food and Drug Administration (FDA) offer only non-binding guidance on fairness and equity in clinical algorithms.[67] These guidelines lack the force of law and are rarely coupled with independent validation or public transparency. As a result, fairness remains optional in the U.S.—subject to internal compliance cultures and voluntary commitments by developers rather than enforceable mandates.

Transparency rights diverge markedly. The GDPR's Articles 13–15 and 22 establish a quasi-"right to explanation," requiring disclosure of algorithmic logic, data provenance, and decision significance in cases of automated processing with substantial effects.[68] This right, though contested in academic literature, has been reinforced by regulatory practice and jurisprudence in the EU, particularly in domains like health and insurance.[69] The U.S., however, lacks any statutory analogue. No federal law mandates disclosure of clinical AI logic to patients or clinicians, and existing medical device regulation does not require transparency in model development or subgroup calibration. This informational asymmetry not only impairs patient autonomy but also undermines the capacity of clinicians to verify model reliability in diverse populations.

The pre-market oversight mechanisms in the EU are designed to be anticipatory and preventive. High-risk clinical AI systems must undergo third-party conformity assessments, including bias audits, traceability checks, and human oversight validation.[70] These audits are conducted by notified bodies empowered to deny certification if subgroup fairness benchmarks are not met. In contrast, the FDA's approach is post hoc and discretionary: while it encourages developers to submit evidence of fairness, there is no enforceable requirement to test model performance across protected demographic groups prior to approval.[71] The absence of mandatory pre-market subgroup validation is particularly concerning given the well-documented performance disparities of machine learning models across race, sex, and socio-economic status.[72]

Liability structures reflect divergent theories of harm and remediation. EU law provides both regulatory enforcement (via national Data Protection Authorities and conformity assessors) and individual remedies, including strict liability for harm caused by high-risk AI systems under the proposed AI Act.[73] This removes the evidentiary burden of proving negligence or discriminatory intent—barriers that often foreclose relief in common law systems. In the U.S., by contrast, plaintiffs must typically rely on tort or civil rights litigation, where establishing causation, discriminatory impact, and developer fault remains a high bar. Courts have struggled to apply traditional liability doctrines to opaque, probabilistic, and distributed algorithmic systems—particularly in healthcare contexts, where harms may manifest subtly or over time.[74]

Finally, only the EU imposes a legal obligation for subgroup validation. The AI Act mandates that developers simulate model outputs across protected categories and include fairness metrics in certification dossiers.[75] This requirement ensures that algorithmic equity is treated not as a downstream quality assurance issue, but as an upstream design criterion. The U.S. lacks any such obligation: no federal statute or FDA rule compels subgroup testing, and developers are not required to disclose whether their models perform unequally across demographic lines. This omission perpetuates the risk of algorithmic harm to already vulnerable populations and limits the regulatory capacity to intervene ex ante.

Together, these differences are not merely procedural—they reflect divergent normative commitments to the role of law in mediating technological risk. The EU treats fairness as a public law value, to be embedded through binding obligations, institutional oversight, and remedial access. The U.S., in contrast, continues to treat algorithmic fairness as an aspirational or ethical concern—delegated to private actors and remedied only when harm can be demonstrated under traditional legal doctrines.

As a result, the EU framework functions as an integrated algorithmic constitutionalism: a layered regime of rights, duties, and procedures that govern the life-cycle of high-risk AI systems in service of democratic values. The U.S. regime, by contrast, operates under what scholars have called "algorithmic laissez-faire"—a patchwork of sectoral rules, soft law instruments, and under-enforced civil rights statutes.[76] For clinical AI systems, which directly affect human health and bodily integrity, this divergence carries urgent implications. Without binding transparency requirements, premarket fairness assessments, and enforceable liability regimes, the U.S. risks institutionalizing structural bias under the guise of technological progress.

Implications for U.S. Legal Reform

Drawing Lessons from the EU Model, U.S. Policymakers—Including the FDA, HHS Office for Civil Rights, and Congress—Should Consider the Following Legislative and Regulatory Interventions:

- **Embed Fairness as a Pre-Market Regulatory Condition:** Like the EU's conformity assessment model, FDA approval processes should require subgroup performance validation across race, gender, and other protected traits as a condition for market authorization.[77]

- **Codify a Statutory Right to Explanation:** Federal legislation—either amending HIPAA or under a new AI law—should guarantee patients and clinicians the right to intelligible, actionable information about algorithmic decision logic and data provenance.

- **Create an Actionable Right Against Algorithmic Discrimination:** Section 1557 should be amended, or new federal legislation introduced, to define algorithmic bias as a form of actionable discrimination, reduce evidentiary burdens, and enable compensatory and injunctive relief.

- **Establish Independent Certification Regimes:** An independent system of third-party AI fairness auditors should

be authorized and accredited to conduct bias impact assessments and fairness validation—replicating the role of EU conformity assessors.

The EU model demonstrates that fairness can be institutionalized effectively through rights-based and regulatory mechanisms." By contrast, U.S. legal frameworks remain disjointed, aspirational, and unable to prevent bias proactively. Incorporating the EU's model of enforceable fairness could significantly enhance public trust and safeguard equitable access to the next generation of AI-enabled clinical care.

A Multi-Layered Fairness Accountability Regime: Designing a Legal-Technical Framework for Algorithmic Equity in U.S. Clinical Decision-Making

The advent and growing integration of algorithmic decision-support tools in clinical settings have sparked both excitement and deep concern. These systems, powered by machine learning and big data analytics, promise improvements in diagnostic accuracy, personalized treatment plans, and optimized resource allocation. Yet, empirical investigations have repeatedly demonstrated that clinical algorithms may replicate or amplify existing health disparities along lines of race, ethnicity, sex, and socio-economic status.[78] The consequences of biased algorithms in healthcare are not merely statistical errors but potential drivers of inequitable treatment outcomes, exacerbating historic structural inequalities.

Despite these pressing challenges, the U.S. regulatory and legal landscape remains fragmented, reactive, and insufficiently equipped to systematically ensure algorithmic fairness in clinical decision-making. The complexity of clinical environments, coupled with the opacity and adaptability of AI systems, demands a multi-dimensional accountability regime that fuses legal standards with technical norms and procedural safeguards.

This section proposes a multi-layered fairness accountability regime that integrates doctrinal reforms, procedural innovations, and technical best practices to establish algorithmic equity as a foundational standard of care in U.S. healthcare. The framework draws on insights from epidemiology, computer science, civil rights law, tort doctrine, administrative law, and comparative international regulatory models. Its design seeks to balance the imperatives of innovation, patient safety, equity, and transparency, thereby advancing a robust legal-technical governance architecture for clinical algorithms.

Algorithmic Fairness as a Standard of Clinical Care

The Clinical Imperative for Subgroup-Specific Fairness

At the heart of algorithmic equity in clinical decision-making lies the recognition that aggregate measures of algorithmic accuracy and performance conceal critical disparities among demographic subgroups. Pioneering work by Obermeyer and colleagues revealed that an algorithm widely used to allocate health resources systematically underestimated the risk for Black patients, leading to their under-referral for care despite higher actual needs.[79] This stark example underscores that fairness must be operationalized through subgroup-specific performance standards, ensuring that predictive accuracy and error rates are equitable for all clinically relevant populations.

Unlike many AI applications where uniform accuracy is desirable, clinical decision tools must account for heterogeneous disease prevalence, physiological differences, and social determinants of health. For instance, an algorithm predicting cardiovascular risk may perform well in the general population but poorly in women or minority groups unless explicitly calibrated and validated for these subpopulations. Thus, setting minimum thresholds for sensitivity, specificity, and predictive values within each subgroup is necessary to uphold the ethical principles of beneficence and justice [22].[80]

Codifying Fairness into the Standard of Care Doctrine

Embedding algorithmic fairness within the legal standard of care in medical malpractice law represents a powerful strategy to incentivize compliance and safeguard patient welfare. The standard of care traditionally reflects the level of diligence and competence expected of a reasonably prudent physician under similar circumstances.[81] Given the increasing reliance on decision-support algorithms, courts should recognize fairness metrics— especially subgroup-specific thresholds—as integral to this evolving standard.

Such recognition would mean that deploying an algorithm failing to meet validated fairness benchmarks could constitute negligence per se or product liability when harm results.[82] This legal development would prompt healthcare providers and developers to rigorously test and audit algorithms on diverse populations, disclosing subgroup performance metrics as part of the clinical validation process. Importantly, the standard of care doctrine's inherent flexibility accommodates continuous updates as new clinical evidence and fairness standards emerge, enabling an adaptive accountability mechanism in a rapidly evolving technological landscape.

Reconceptualising Anti-Discrimination Law for Algorithmic Fairness in U.S. Clinical Decision-Making

As algorithmic tools increasingly inform clinical decision-making, the foundational legal architecture of anti-discrimination law must evolve to address novel patterns of harm. Algorithmic systems—particularly those embedded in clinical decision support software (CDSS)—introduce complex, datadriven forms of bias that may operate independently of human intent yet result in racially or demographically disparate outcomes. Traditional antidiscrimination law, oriented around intentionality and overt exclusion, is ill-equipped to redress such harms. In this section, we propose a retooling of anti-

discrimination doctrine to accommodate the epistemic and structural realities of algorithmic bias, with a specific focus on the underutilised potential of the Affordable Care Act's Section 1557.

Section 1557 ACA as a Normative and Regulatory Lever

Section 1557 of the Patient Protection and Affordable Care Act (ACA) prohibits discrimination "on the ground[s] prohibited under" Title VI of the Civil Rights Act of 1964, Title IX of the Education Amendments Act of 1972, the Age Discrimination Act of 1975, and Section 504 of the Rehabilitation Act of 1973 in "any health program or activity" receiving federal financial assistance.[83] It incorporates a broad mandate of nondiscrimination across race, colour, national origin, sex, age, and disability, and has been interpreted to cover both public and private health institutions receiving federal funds.

However, its application to algorithmic discrimination, particularly in the absence of direct human intent, remains underdeveloped. Unlike Title VII jurisprudence—which recognises disparate impact liability under Griggs v. Duke Power Co.[84]—Section 1557's treatment of indirect discrimination by AI remains ambiguous. Courts have yet to definitively address whether outcomes produced by machine learning systems, trained on historically biased datasets, fall within the scope of actionable harm under this statute.

To bridge this doctrinal lacuna, federal regulatory bodies such as the Department of Health and Human Services (HHS) should issue interpretive rules or formal guidance under Chevron-deference principles[85] affirming that algorithmic disparate impacts on protected classes violate Section 1557, regardless of developer intent. This approach finds analogical support in the Department of Education's Title VI disparate impact regulations[86] and recent efforts by the Equal Employment Opportunity Commission (EEOC) to apply impact-based liability in algorithmic hiring practices.[87]

Crucially, the proposed interpretation would enable pre-emptive regulation, shifting legal scrutiny to outcomes rather than the subjective mental states of clinicians, developers, or institutions. Such a shift is essential in light of empirical findings demonstrating that algorithmic systems can consistently under-predict risk for Black patients in resource allocation tools, leading to large-scale inequities in care delivery.[88]

Embedding Procedural Safeguards: Algorithmic Due Process and Auditability

The procedural invisibility of machine learning models—often described as operating within a "black box" —complicates traditional mechanisms of anti-discrimination enforcement. Without transparent or explainable system architectures, injured parties struggle to prove causation, disparate impact, or foreseeability, which are prerequisites for both tort and statutory redress. Accordingly, to operationalise Section 1557 in the algorithmic era, procedural safeguards must be institutionalised as both regulatory and constitutional obligations.

We propose the mandatory adoption of algorithmic fairness procedures analogous to administrative due process principles established in Mathews v. Eldridge, which require balancing the private interest affected, the risk of erroneous deprivation, and the government's interest in streamlined adjudication.[90] In algorithmic healthcare contexts, the stakes are often life-altering, and the opacity of the decision logic renders informal protections insufficient.

Three Safeguards are Foundational:

- **Independent Bias Audits:** Clinical algorithms must be subject to third-party audits conducted by accredited fairness assessors. Audits should include statistical parity testing, subgroup performance stratification, and calibration curves disaggregated by protected characteristics. Findings must be publicly reported to foster transparency and downstream accountability.[91]

- **Disclosure of Subgroup Performance Metrics:** As part of regulatory filings and institutional compliance reviews, developers and healthcare providers should be compelled to disclose granular performance metrics across race, gender, age, disability status, and socio-economic background. This parallels emerging obligations under the EU AI Act and reinforces patient-centred transparency norms.[92]

- **Explainability Standards:** Systems deployed in clinical settings must adhere to a minimum threshold of interpretability, using methods such as LIME (Local Interpretable Model-Agnostic Explanations), SHAP (Shapley Additive Explanations), or transparent surrogate models.[93] These methods enable clinicians and patients to interrogate outputs and challenge decisions, enhancing both informed consent and legal contestability.

These procedural mechanisms, if embedded within the enforcement apparatus of Section 1557, would convert an often inert anti-discrimination framework into an anticipatory regulatory tool. They operationalise constitutional values of fairness and individual dignity while aligning with administrative principles of transparency, accountability, and reviewability.

Towards a Rights-Based Model of Algorithmic Health Equity

In aligning Section 1557 with the demands of algorithmic equity, a broader jurisprudential shift is required: from a model of discrimination focused solely on animus and intent to one grounded in systemic outcome responsibility. This mirrors developments in international human rights law, where states are increasingly held accountable for indirect discrimination

arising from ostensibly neutral technologies.[94] The European Court of Human Rights, for instance, has recognised indirect discrimination where facially neutral practices have disproportionate impacts, particularly in healthcare access.[95]

This global rights-based evolution offers a template for U.S. courts and regulators. By acknowledging algorithmic bias as a structural rights violation, rather than a mere artefact of data error or development oversight, Section 1557 can be deployed as a transformative tool. Courts should interpret the statute not as a passive barrier against overt discrimination but as an affirmative mechanism to scrutinise, rectify, and prevent technological reproduction of health disparities.

Furthermore, recognition of algorithmic harm under Section 1557 would facilitate regulatory harmonisation with other federal mandates, such as the 21st Century Cures Act, which mandates transparency in clinical decision support software, and the Health Information Technology for Economic and Clinical Health (HITECH) Act, which promotes patient access to digital health records.[96]

As clinical algorithms become integral to diagnostic and treatment workflows, failure to adapt anti-discrimination frameworks risks entrenching racial and socio-economic disparities under a veneer of technocratic objectivity. Section 1557, though historically under-applied in this context, offers a doctrinally sound and normatively compelling foundation for algorithmic equity. By embracing disparate impact theory, embedding procedural safeguards, and expanding conceptions of discrimination to account for algorithmic systems, U.S. health law can align itself with both global norms and ethical imperatives. Such reform is not merely doctrinal—it is infrastructural. It reimagines the state's role not only as a referee of past wrongs but as a steward of future fairness, capable of shaping a digital health ecosystem that is inclusive, accountable, and just.

Designing a No-Fault Algorithmic Injury Compensation Scheme: Towards Equitable Redress in the Age of Clinical AI

Redressing Structural Injustice Beyond Tort: The Case for an Algorithmic Compensation Fund

As artificial intelligence becomes increasingly embedded in clinical decision-making, the legal system faces a critical legitimacy challenge: how to ensure just and timely compensation for patients harmed by algorithmic errors that are probabilistic, emergent, and often opaque.[97] Even with improved regulatory oversight and enhanced civil rights enforcement, the inherent uncertainty and adaptive nature of machine learning-based clinical decision support systems (AI-CDSS) mean that adverse outcomes—including racially skewed recommendations, false negatives, and inappropriate triage—are not only possible but statistically inevitable.[98]

Traditional tort litigation remains structurally misaligned with the contours of these harms. Doctrines such as negligence and product liability hinge on the identification of a discrete act of wrongdoing, typically grounded in human fault or product defectiveness at a static point in time.[99] However, AI systems evolve dynamically, interact with heterogeneous clinical contexts, and frequently involve distributed responsibility across multiple actors— developers, clinicians, data scientists, and healthcare institutions.[100] Moreover, litigation imposes substantial procedural and financial burdens on plaintiffs, disproportionately affecting socio-economically disadvantaged groups who are most likely to suffer from embedded algorithmic bias.[101]

In this light, a no-fault algorithmic injury compensation fund offers a normatively attractive and practically efficient mechanism for providing redress. Modelled on established administrative regimes such as the National Vaccine Injury Compensation Program (VICP)[102] and the Countermeasures Injury Compensation Program (CICP),[103] such a fund would enable patients harmed by AI-driven clinical tools to obtain compensation without proving intent, negligence, or product defect. It would recognize algorithmic harm as a systemic externality rather than an individualized failure—thus shifting the burden of injury resolution from adversarial litigation to collective accountability.

From a jurisprudential perspective, this proposal aligns with foundational principles of distributive justice and corrective equity, acknowledging that algorithmic systems are embedded within complex institutional and sociotechnical ecologies.[104] The fund would not replace tort entirely but would instead serve as a supplemental, low-threshold alternative capable of addressing diffuse harm at scale.

Structure, Governance, and Funding Mechanisms

To ensure legitimacy, efficiency, and equity, the algorithmic compensation fund must be carefully structured along principles of expert adjudication, transparent eligibility criteria, and multi-stakeholder financing.

- **Administration and Adjudication:** The fund could be administered by the Department of Health and Human Services (HHS) or a newly constituted Algorithmic Injury Compensation Authority (AICA). Claims would be adjudicated by interdisciplinary expert panels comprising clinicians, ethicists, data scientists, legal scholars, and patient advocates. These panels would assess causation and eligibility based on pre-established risk profiles, audit trails, and subgroup performance data. This adjudicatory model draws on the expertise-based determination structures used in the VICP and offers epistemic legitimacy absent in traditional jury-based tort systems [23].[105]

- **Eligibility Criteria and Burden of Proof:** Claimants would not be required to prove negligence or intentionality.

Instead, they would demonstrate that an adverse outcome occurred and that the deployed AI-CDSS fell within the fund's list of covered tools and known risk profiles. Eligibility would also take into account patterns of bias identified through post-deployment algorithmic audits, model drift analyses, and real-world performance disparities documented by the FDA or clinical institutions.[106]

• **Funding Architecture:** The scheme would be financed through mandatory contributions from stakeholders who profit from or deploy algorithmic tools, including: AI developers and vendors (via licensing or per-unit fees); healthcare institutions deploying clinical AI systems (via premium assessments); and insurers underwriting algorithm-integrated care pathways.

This cost-internalisation mechanism mirrors that of the VICP, which is funded via excise taxes on each vaccine dose.[107] Crucially, it promotes ex ante investment in fairness and safety by embedding risk-adjusted pricing into the development and deployment pipeline. To further incentivise compliance, participation in the fund could be made a condition of FDA approval or CMS reimbursement eligibility.[108] Such a regulatory lever would encourage widespread adoption and signal systemic commitment to ethical AI governance.

Comparative Models and Global Precedents

Internationally, several jurisdictions are experimenting with no-fault schemes for digital health innovation. For example, the United Kingdom's NHS Resolution is exploring the expansion of its indemnity model to encompass AI-related harms, while Canada's Compensation for Vaccine Injuries Program demonstrates the administrative feasibility of low-threshold claims adjudication for medical technologies.[109] The logic of algorithmic compensation also echoes recent proposals from the European Parliament, which advocates for a harmonised civil liability framework to address AIspecific risks through strict liability and mandatory insurance schemes.[110] While these models differ in legal design, they converge on a shared recognition: that digital systems require institutional infrastructures of accountability beyond case-by-case tort adjudication.

In adopting a U.S.-specific model, policymakers must balance federalism concerns, stakeholder resistance, and potential moral hazard. However, as algorithmic tools become more deeply embedded in Medicare and Medicaid-funded services, federal jurisdiction and public interest clearly support the creation of a centralized, equity-oriented redress mechanism.

A no-fault algorithmic injury compensation fund represents more than a procedural innovation—it is a normative reimagining of how law responds to technological risk. By shifting from individualised fault to systemic stewardship, such a fund can embody the ethical commitment to patient safety, distributive justice, and democratic oversight in the governance of clinical AI. It provides a crucial component of a broader regulatory framework—one that ensures algorithmic medicine serves, rather than undermines, the foundational values of equity and dignity in healthcare delivery.

Integrated Governance: Toward a Coordinated Regulatory Architecture for Algorithmic Fairness

Fragmented Oversight in a Converging Technological Landscape

Artificial intelligence (AI) in clinical decision-making occupies an increasingly complex and fragmented regulatory space. In the United States, disparate agencies regulate discrete components of the AI life-cycle: the Food and Drug Administration (FDA) governs software as a medical device (SaMD); the Office for Civil Rights (OCR) within the Department of Health and Human Services (HHS) enforces anti-discrimination laws under Section 1557 of the Affordable Care Act; the Centres for Medicare & Medicaid Services (CMS) control reimbursement and value-based care incentives; and the Federal Trade Commission (FTC) exercises jurisdiction over data privacy and deceptive practices [24].[111]

This diffusion of authority, although reflecting the multifaceted nature of clinical AI systems, creates significant risks of regulatory lacunae, overlapping mandates, and inconsistent enforcement.[112] Divergences in definitions (e.g., what constitutes "bias," "fairness," or "meaningful transparency"), legal standards, and thresholds for intervention compound these gaps, especially when algorithms intersect civil rights, medical safety, and data governance simultaneously.[113] The consequences are not merely theoretical: multiple clinical AI tools have entered the healthcare system with insufficient subgroup validation, opaque performance metrics, and disparate impact on vulnerable populations—despite passing regulatory muster under one domain.[114]

To address these limitations, an inter-agency governance model is essential. This model should take the form of a permanent, cross-sectoral Algorithmic Health Oversight Task Force housed within HHS or the Executive Office of the President, charged with harmonising legal standards and coordinating enforcement across federal regulators.[115] This proposal builds on the structure of the National AI Initiative Office and recent Biden Administration directives promoting cross-agency AI coordination, including Executive Order 14110 on Safe, Secure, and Trustworthy AI.[116]

Critically, this task force must integrate civil rights, health equity, clinical efficacy, and technical robustness as coequal pillars of regulatory concern— not siloed domains. It should function through joint rulemaking authority, pooled investigatory resources, and a centralised clearing-house for AIrelated complaints, audits, and enforcement referrals.[117] By institutionalising coordination rather than relying on ad hoc agency collaboration, the Task Force would operationalise the principle that algorithmic fairness is not a niche ethical concern but a structural regulatory imperative.

Dynamic Standard-Setting and Responsive Enforcement

Traditional regulatory models are ill-suited to the fast-paced evolution of machine learning (ML) technologies, whose outputs are probabilistic, nondeterministic, and context-dependent. To mitigate this mismatch, regulatory standards must be dynamic, iterative, and risk-based, capable of evolving with real-world evidence, clinical feedback, and algorithmic drift.

The FDA's Digital Health Software Precertification (Pre-Cert) Program represents an early prototype of such adaptive regulation. Though the program is currently paused for reassessment, it introduced the principle of total product life-cycle (TPLC) oversight—emphasising continuous performance monitoring rather than static pre-market clearance.[118] This model is especially relevant for AI-CDSS tools that adapt over time, either via federated learning, periodic retraining, or environment-sensitive updates.[119]

To Embed Fairness Within this Adaptive Regulatory Paradigm, the Following Components Should be Mandated:

- **Pre-Market Fairness Validation:** Clinical AI tools must undergo subgroup performance analysis and algorithmic impact assessments prior to deployment, with results publicly disclosed and certified.

- **Post-Market Surveillance for Bias:** Regulators should require continuous real-world monitoring of algorithmic outputs for disparate impact using equity-sensitive quality measures, such as those developed by the Agency for Healthcare Research and Quality (AHRQ).[120]

- **Corrective Action Mandates:** If post-market data reveals unjustifiable disparities or harms, regulators must possess robust recall powers, penalty provisions, and injunctive remedies tailored to algorithmic tools.[121]

- **Administrative Penalties and Transparency Incentives:** Developers failing to comply with fairness mandates should face civil monetary penalties, while those exceeding transparency standards could qualify for expedited regulatory pathways or CMS reimbursement bonuses.[122]

Moreover, an integrated governance system must require algorithmic explainability as a procedural due process right. Echoing jurisprudence from Mathews v. Eldridge,[123] this would ensure that affected individuals, especially in high-stakes contexts such as organ allocation or oncology, are not subject to black-box decisions devoid of contestability.[124]

The regulatory regime must not only ensure ex post redress but also generate ex ante incentives for fairness, safety, and transparency across the algorithmic pipeline. In line with the theory of responsive regulation, soft guidance (e.g., fairness toolkits, transparency templates, and validation checklists) should be coupled with hard enforcement (e.g., penalties, debarment, and litigation referrals).[125] Publicly accessible registries of certified clinical algorithms, complete with subgroup performance data, would bolster market accountability and patient trust.

Clinical AI governance cannot succeed through technical fixes or siloed intervention. The structural embedding of fairness in algorithmic healthcare requires a paradigm shift: from reactive, fragmented oversight to coordinated, anticipatory, and equity-oriented regulation. An integrated interagency framework—grounded in shared principles, responsive standard-setting, and robust enforcement—offers the clearest path to ensure that clinical AI systems not only perform effectively but also advance the public interest, protect vulnerable populations, and uphold the constitutional values that animate the U.S. healthcare system.

The Multi-Dimensional Fairness Matrix: Aligning Technical Metrics, Legal Doctrines, and Clinical Ethics

To render algorithmic fairness both actionable and enforceable in U.S. clinical decision-making, a multi-dimensional fairness matrix must be institutionalized—one that operationalizes fairness across three interlocking axes: computational performance metrics, foundational legal principles, and core clinical ethical values. The matrix presented below (Figure 3) exemplifies this triangulation, translating abstract notions of equity into governance-ready, context-specific benchmarks.

Fairness Metric	Legal Principle	Clinical Ethical Value
Equalized Odds (balanced error rates)	Anti-Discrimination Law	Justice and Non-Maleficence
Subgroup Sensitivity and Specificity Thresholds	Standard of Care in Tort Law	Beneficence and Equity
Transparency and Explainability	Procedural Fairness (Due Process)	Autonomy and Accountability
Error Minimization and Harm Reduction	Liability and Compensation	Safety and Trust

Figure 3: The Multi-Dimensional Fairness Matrix

From Statistical Fairness to Regulatory Norms: Translating Metrics into Legal Doctrine

Statistical fairness metrics such as equalized odds, predictive parity, and calibration are indispensable to identifying algorithmic disparities across racial, gender, age, and socio-economic subgroups. Yet their technical precision belies an unresolved tension: these metrics often trade off against each other and may inadvertently entrench structural inequities when applied without normative grounding.[126] For example, optimizing for predictive parity (ensuring equal

positive predictive value across groups) can exacerbate disparities in false negatives for under-diagnosed populations—a critical issue in oncology, cardiology, and mental health AI tools.[127]

This is where legal theory must act as an interpretive scaffold. Equalized odds, which demand parity in false positive and false negative rates across protected subgroups, aligns most directly with anti-discrimination norms under U.S. constitutional and statutory law, including Title VI of the Civil Rights Act and Section 1557 of the Affordable Care Act.[128] Courts have historically embraced disparate impact frameworks under these statutes, even when intent to discriminate is not provable—allowing statistical disparities to form the evidentiary basis for liability.[129] Thus, codifying equalized odds into pre-market approval standards offers a legally coherent and ethically sound pathway to embedding fairness into regulatory design.

Subgroup Thresholds and the Standard of Care in Tort Law

Subgroup-specific sensitivity and specificity thresholds—measures of a model's accuracy within demographic subsets—are crucial to ensuring that no patient group systematically receives inferior care due to algorithmic design flaws. These thresholds mirror tort law's evolving "standard of care" doctrine. When deployed in clinical settings, AI tools can influence, reinforce, or even supplant clinical judgment, rendering their outputs material to malpractice adjudication.

Under current doctrine, liability hinges on whether a clinician's reliance on an AI recommendation breaches a duty of reasonable care.[130] However, absent established technical performance standards, courts lack consistent benchmarks. Mandating demographic performance floors provides a judicially cognizable standard for both clinicians and AI developers, much like clinical guidelines shape standard-of-care determinations in diagnosis or treatment.

Transparency, Explainability, and Procedural Due Process

Transparency and explainability—core tenets of algorithmic accountability—resonate deeply with administrative and constitutional commitments to procedural fairness. AI tools that are inscrutable to end-users (clinicians) and stakeholders (patients) challenge the core values underpinning due process under the Fifth and Fourteenth Amendments.

The U.S. Supreme Court's decision in Mathews v. Eldridge set out a balancing test for procedural due process that hinges on the risk of erroneous deprivation, the value of additional safeguards, and the government's interest in administrative efficiency.[131] Applied to algorithmic medicine, explainability becomes essential not only for clinical trust, but also for ensuring that patients can meaningfully contest outcomes and seek recourse. These mechanisms—when embedded into FDA review processes or required under Section 1557—reaffirm the right of patients to understand and challenge decisions materially affecting their health.

Minimizing Algorithmic Harm Through Liability and Compensation

The principle of non-maleficence—the ethical obligation to avoid harm—demands not only risk mitigation ex ante but effective remedies ex post. In algorithmic contexts, this translates to robust error minimization protocols (such as fail-safe systems and model degradation monitors) and responsive compensation mechanisms.

In tort law, the increasing use of probabilistic tools complicates causation analyses and places significant evidentiary burdens on harmed patients, particularly when the harm stems from latent design bias rather than discrete clinical acts. A no-fault algorithmic injury compensation fund, as discussed in Section 5.3, not only aligns with principles of restorative justice and distributive equity but also ensures systemic resilience. It creates incentives for developers to invest in safer design while providing patients with a predictable path to redress.

Clinical Ethics as a Guiding Framework: From Norms to Design Choices

Each of the above technical-legal intersections is undergirded by bioethical values historically central to U.S. clinical governance. Justice, nonmaleficence, beneficence, autonomy, and accountability are not abstract ideals; they are actionable imperatives that guide everything from informed consent to institutional review board (IRB) protocols.[132]

Integrating these values into algorithmic design requires more than post hoc validation—it requires participatory design, cross-disciplinary deliberation, and clinical interpretability at the point of care. For instance, models developed without input from historically marginalized patient communities may reproduce the very inequities fairness metrics aim to redress. A fairness framework that fails to incorporate clinical ethics risks technical robustness without moral legitimacy.

Operationalizing the Matrix: A Regulatory Blueprint

To Move From Conceptual Matrix to Enforceable Regime, Several Institutional Strategies are Necessary:

- **Codification into Regulatory Guidance:** The FDA, HHS OCR, and CMS must issue joint interpretive rules aligning technical metrics with civil rights and reimbursement compliance frameworks.

- **Standardization Through Public-Private Collaboration:** Federal coordination with clinical accreditation bodies (e.g. The Joint Commission), AI developers, and professional societies can help align technical standards with practice norms.

• **Integration into Clinical Trials:** Fairness metrics and explainability audits should be integrated into pivotal clinical trials for AI-CDSS approval—much like subgroup analyses are now standard in drug trials.[133]

• **Public Reporting and Equity Dashboards:** Publicly accessible registries of AI-CDSS tools, including subgroup performance data, bias mitigation plans, and post-market surveillance outcomes, can enhance transparency and drive continuous quality improvement.

The multi-dimensional fairness matrix offers a rigorous, interdisciplinary blueprint for aligning algorithmic performance metrics with legal accountability and clinical ethical obligations. By institutionalizing this triadic framework, regulators, clinicians, developers, and courts can converge on a shared language of fairness—one that is not only measurable but enforceable and morally justified. In a domain where the stakes are no less than life and death, such integration is not aspirational—it is imperative.

Comparative and International Lessons for U.S. Algorithmic Fairness in Clinical AI

The regulation of algorithmic fairness in clinical decision-making is not solely a domestic concern but a global policy imperative. Comparative analysis reveals that several jurisdictions—most notably the European Union and the United Kingdom—have already moved to embed enforceable fairness and transparency obligations within their legal and regulatory ecosystems. For the U.S., this international momentum presents both a normative benchmark and an opportunity: to develop a uniquely American approach that fuses technical excellence with civil rights accountability, while learning from international best practices.

The European Union's AI Act: Risk-Based Regulation and Embedded Fairness

The European Union's draft Artificial Intelligence Act (AI Act) represents a paradigmatic shift in regulatory philosophy: a move from reactive enforcement toward proactive, ex ante risk governance. It classifies AI systems into four risk tiers—unacceptable, high, limited, and minimal—based on the potential impact on fundamental rights, safety, and social outcomes.[134] Clinical decision support systems (CDSS), particularly those involved in diagnosis or treatment prioritization, are generally categorized as high-risk, thereby subject to strict oversight obligations.

Key provisions of the AI Act include: Mandatory bias and discrimination risk assessments during development, traceability and documentation obligations across the life-cycle of the model (Article 10), human oversight requirements to ensure that clinicians can contest or override AI outputs (Article 14), independent third-party conformity assessments prior to deployment (Article 43), post-market monitoring obligations and reporting of serious incidents (Articles 61–62).

These mechanisms function as fairness enforcement tools. By legally mandating documentation of data provenance, subgroup performance, and mitigation strategies, the EU imposes a regulatory duty of care that both anticipates harm and centres patient rights as part of algorithmic governance. Unlike the U.S., where algorithmic bias is often addressed through retrospective litigation, the EU framework integrates fairness into the design, deployment, and post-deployment stages, aligning with principles of precautionary regulation and fundamental rights by design.[135]

The United Kingdom's NHS Algorithmic Transparency Standard: Procedural Fairness and Stakeholder Engagement

The UK's NHSX (now under the Department of Health and Social Care) has pioneered a complementary yet distinctive approach through its "Algorithmic Impact Assessment" (AIA) and "Transparency Standard" for AI in healthcare. While non-binding, these instruments emphasize procedural justice, inclusive design, and public accountability.[136]

Notably, the NHS framework requires developers to: Disclose the purpose, data sources, and expected outputs of clinical algorithms; publish fairness audits and impact assessments for public review; engage diverse stakeholders—including clinicians, ethicists, and patients—at early design stages; and undergo external validation and continuous monitoring through feedback loops.

Unlike the EU's prescriptive legal obligations, the NHSX framework leans on soft law instruments and institutional norms. However, its emphasis on deliberative ethics, public engagement, and institutional transparency provides a critical complement to the EU's more technocratic risk calculus. From a comparative legal perspective, the UK's approach reflects a procedural fairness model, rooted in the legacy of administrative law and rights-based health governance under the National Health Service Act 2006.

Lessons for U.S. Regulatory Adaptation: Synthesis, Civil Rights, and Systemic Equity

The U.S. lacks a unified regulatory framework for AI in healthcare; instead, jurisdiction is divided among the Food and Drug Administration (FDA), Office for Civil Rights (OCR) under HHS, Centres for Medicare & Medicaid Services (CMS), and the Federal Trade Commission (FTC)—each focusing on different dimensions of safety, privacy, access, or consumer protection. While this institutional dispersion enables issue-specific expertise, it also creates regulatory gaps where no single agency is accountable for algorithmic fairness in clinical outcomes.

What the U.S. can Adapt—and Arguably Improve Upon—Includes

- **Embedding Fairness into FDA Premarket Review** through enforceable subgroup performance thresholds and explainability audits, much like the EU's conformity assessments;

- **Mandating Public Algorithmic Impact Assessments** as a condition for reimbursement under CMS, similar to the NHS's transparency standard;

- **Integrating Title VI and Section 1557 Enforcement** within algorithm deployment, thereby making fairness a civil rights compliance requirement, not merely a safety consideration;

- **Facilitating Public Registries and Post-Market Reporting,** akin to both the EU and UK models, to enable real-time monitoring and patient empowerment;

- **Creating a Centralized AI Fairness Coordination Body,** possibly through an Executive Order or congressional mandate, to harmonize interagency standards.

Such a hybridized U.S. model would not be a simple import of European or British norms. Rather, it would leverage the unique legal infrastructure of U.S. civil rights law, particularly the robust doctrines of disparate impact, equal protection, and due process, to enforce algorithmic fairness not as a technological nicety, but as a legal obligation anchored in democratic accountability.

The U.S. is at a critical juncture: either perpetuate fragmented, reactive responses to algorithmic bias in clinical settings or draw from emerging global frameworks to institutionalize a robust, enforceable fairness regime. The EU's AI Act and the UK's NHSX standards offer divergent yet complementary models—one grounded in binding risk regulation, the other in procedural and democratic legitimacy. For the U.S., the optimal path forward lies in synthesizing these insights through the lens of its constitutional commitments, federalist structures, and moral leadership in civil rights. Such an approach would not only improve domestic health equity but position the U.S. as a global norm-setter in ethical AI governance.

Conclusion

The proliferation of machine learning and artificial intelligence (AI) in clinical decision-making systems presents a profound regulatory challenge—one that intersects law, technology, bioethics, and public health. Algorithmic bias, if left unchecked, risks reinforcing and amplifying pre-existing health inequities, particularly across racial, socio-economic, gender, and disability-based lines. Consequently, this section has outlined a normative and regulatory roadmap for constructing a multi-dimensional fairness accountability regime in the U.S., rooted in legal rigour, technical sophistication, and ethical integrity.

At its foundation, such a framework must reconfigure the standard of care in medical tort law to include subgroup-specific fairness obligations—an evolution consistent with the growing judicial receptivity to data-driven standards in Helling v Carey[137] and Canterbury v Spence[138], where clinical norms evolve with medical knowledge and technological change. AI tools that exhibit clinically significant disparate error rates across protected groups must no longer be considered reasonable or customary if they fail fairness validation benchmarks—especially when safer, fairer alternatives exist.

Simultaneously, robust civil rights enforcement under Section 1557 of the Affordable Care Act and Title VI of the Civil Rights Act must be expanded to encompass algorithmic discrimination in clinical contexts. Courts and agencies must clarify that disparate impact in clinical AI outputs—where foreseeable, preventable, and rooted in data or model design—constitutes a legally actionable harm, drawing on the interpretive reasoning in Alexander v Sandoval[139] and Griggs v Duke Power Co[140].

Critically, the adoption of pre-market subgroup validation mandates, modelled after EU-style conformity assessments under the Artificial Intelligence Act[141], would embed fairness into clinical algorithm design rather than retroactively chasing harm. These obligations must be backed by administrative enforcement powers, civil penalties, and—where appropriate—no-fault compensation schemes that provide remedies without requiring proof of negligence or intent, as seen in comparative healthcare models like Sweden's Patient Injury Act and New Zealand's Accident Compensation Corporation.

To ensure coherence across fragmented federal authorities, a national inter-agency oversight architecture—possibly housed within the Department of Health and Human Services (HHS)—is necessary. This body would coordinate the FDA, OCR, CMS, and FTC in aligning risk classification, fairness thresholds, bias mitigation protocols, and data governance standards. Inspiration may be drawn from the U.S. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (2023), which calls for cross-sector algorithmic accountability frameworks. Finally, the proposed multi-dimensional fairness matrix—aligning technical fairness metrics (e.g., equalized odds, subgroup sensitivity), legal doctrines (e.g., antidiscrimination law, due process), and clinical ethical imperatives (e.g., justice,

non-maleficence, beneficence)—offers a practical schema to guide developers, regulators, and clinicians in navigating inevitable trade-offs.

Incorporating insights from global regulatory experiments—particularly the EU AI Act, the UK's NHS Algorithmic Transparency Standard, and the WHO's Ethics and Governance of Artificial Intelligence for Health—the U.S. can craft a hybrid model that balances innovation with equity, autonomy with accountability, and precision with protection. In doing so, it would not only mitigate the systemic risks of algorithmic injustice in medicine but also position itself as a global leader in shaping the future of ethical, lawful, and trustworthy clinical AI.

Conclusions and Policy Recommendations: Toward Enforceable Algorithmic Fairness in U.S. Healthcare and Beyond

The integration of Big Data analytics and artificial intelligence (AI) into clinical decision support systems (CDSS) is reshaping the architecture of modern healthcare, heralding advancements in diagnostic precision, therapeutic optimization, and population-level health management. Yet, as demonstrated throughout this study, these technological gains remain precariously tethered to systemic vulnerabilities when deployed without enforceable regulatory guardrails. Empirical evidence confirms that algorithmic models—trained on historically biased datasets and optimized without subgroup fairness metrics—can perpetuate and exacerbate structural inequities across race, gender, socio-economic status, and disability.[142]

This article has exposed how existing U.S. regulatory instruments—spanning FDA oversight of software as a medical device (SaMD), antidiscrimination enforcement under Section 1557 of the Affordable Care Act, and fragmented agency jurisdiction—remain ill-equipped to address the compound risks of algorithmic discrimination in clinical settings.[143] In the absence of a coherent legal-technical framework, such systems threaten to entrench health disparities under a veneer of scientific objectivity and computational neutrality.[144]

Section VI offers a roadmap for reform grounded in doctrinal innovation, technical standardisation, and ethical governance. It proposes a multi-layered accountability architecture—spanning pre-market fairness validation, expanded civil rights enforcement, adaptive inter-agency oversight, and crosssectoral legal harmonisation. Drawing on comparative insights from the EU Artificial Intelligence Act,[145] the UK NHS Algorithmic Transparency Standard,[146] and the WHO's ethical guidance on AI in health,[147] this section articulates how the U.S. can operationalise enforceable algorithmic fairness in healthcare. Beyond the domestic scope, these policy recommendations carry broader normative implications for global digital health governance in an era defined by algorithmic mediation and biomedical datafication.

Summary of Core Findings

Our interdisciplinary analysis reveals substantial legal, technical, and institutional deficiencies in the current U.S. regulatory landscape governing algorithmic fairness in clinical artificial intelligence (AI). These deficiencies arise from the fragmented interaction of medical device regulation, civil rights law, tort doctrine, and health information governance—none of which, individually or collectively, establish enforceable safeguards against algorithmic discrimination in clinical contexts.

FDA Oversight Remains Technically Narrow and Legally Under-Inclusive

While the U.S. Food and Drug Administration (FDA) regulates AI-based clinical decision support tools as Software as a Medical Device (SaMD), its current mandate under the Food, Drug, and Cosmetic Act (FDCA) emphasizes safety and efficacy without mandating pre-market fairness audits or post-market subgroup performance reporting [25].[148] There is no binding requirement for developers to demonstrate parity in error rates across protected sub-populations, despite evidence that such disparities may lead to systemic under-diagnosis in racial minorities.[149] Unlike the European Union's AI Act, which categorizes many clinical algorithms as "high-risk" and subjects them to conformity assessments including fairness testing,[150] the FDA lacks a statutory mechanism to enforce equity-driven algorithmic evaluation.

Civil Rights Protections Under Section 1557 of the ACA are Incomplete

Section 1557 of the Affordable Care Act, while prohibiting discrimination on the basis of race, sex, disability, and national origin in federally funded healthcare programs,[151] has not been judicially extended to encompass disparate impact claims involving algorithmic outputs. Courts remain reluctant to apply traditional Title VI disparate impact reasoning to complex AI systems, in part due to doctrinal uncertainty regarding foreseeability, causation, and the opacity of algorithmic decision-making.[152] Furthermore, administrative enforcement by the Office for Civil Rights (OCR) has been slow to adapt to the algorithmic turn in healthcare delivery, despite recent regulatory efforts to address digital bias [26].[153]

Medical Tort Law Lacks Causal Granularity for Algorithmic Harms

Traditional doctrines of medical malpractice and product liability are ill-equipped to resolve harms arising from biased AI systems, particularly where error differentials are rooted in biased training data rather than identifiable clinical negligence or product defects [27].[154] The doctrinal requirement of individualized fault is misaligned with the systemic nature of algorithmic harm, which is often statistical, probabilistic, and distributed across populations. Without recalibrating the standard of care to incorporate subgroup fairness obligations—aligned with evolving jurisprudence in Helling v Carey[155]

and Canterbury v Spence[156]—courts may continue to defer to biased clinical norms as "reasonable," thereby immunizing discriminatory technologies.

Comparative Jurisdictions Provide Actionable Regulatory Models

The European Union's General Data Protection Regulation (GDPR), particularly the right to meaningful explanation under Article 22,[157] and the proposed AI Act, offer instructive counterpoints. The AI Act introduces binding ex ante obligations, including fairness risk classification, mandatory bias mitigation strategies, and conformity assessments for high-risk applications—most notably in healthcare.[158] Similarly, Sweden and New Zealand have implemented no-fault compensation regimes that allow for equitable redress without proving negligence, thereby bridging the accountability gap in health-tech harms.[159] The U.S., by contrast, remains without a unified regulatory regime capable of addressing these challenges at scale.

Taken together, these findings underscore the urgency for comprehensive regulatory reform that harmonizes doctrinal, ethical, and technical dimensions of algorithmic fairness. A coherent legal regime should mandate independent fairness audits, require clinical AI systems to meet subgroup performance thresholds, and empower courts and agencies to treat biased algorithms as a legally cognizable harm. Absent these changes, clinical AI risks not only entrenching health inequities but also eroding the legal and ethical foundations of the U.S. healthcare system.

Policy Recommendations for U.S. Regulators

In light of the doctrinal, regulatory, and empirical deficits identified across Sections I–V, this section proposes a series of policy interventions to establish a robust, enforceable framework for algorithmic fairness in U.S. clinical decision-making. These recommendations are grounded in comparative legal analysis, normative theory, regulatory science, and empirical scholarship. They target key institutional levers within the Food and Drug Administration (FDA), Department of Health and Human Services (HHS), and Congress to ensure that algorithmic tools advance—rather than undermine—health equity, safety, and democratic accountability.

Empowering the FDA with a Statutory Mandate for Algorithmic Fairness

To address the structural absence of fairness oversight within the FDA's current remit, Congress should amend the Federal Food, Drug, and Cosmetic Act (FDCA) to require algorithmic subgroup fairness as a precondition for pre-market approval of AI/ML-based medical devices [28].[160] These fairness requirements must extend beyond general efficacy to encompass parity in performance across protected characteristics—such as race, ethnicity, sex, disability status, and age—evaluated through disaggregated metrics including subgroup sensitivity, specificity, calibration, false positive/negative rates, and predictive parity [29].[161] This aligns with calls from bioethics and clinical AI scholarship to integrate distributive justice principles into regulatory science [30].[162]

Moreover, the FDA should establish a standing AI Fairness Advisory Committee, composed of interdisciplinary experts in algorithm auditing, medicine, law, data science, and public health. This body would be tasked with updating fairness guidelines, establishing benchmark datasets for fairness testing, and overseeing post-market surveillance to monitor algorithmic drift and real-world subgroup disparities.[163] Such oversight parallels the model of continuous post-market surveillance under the European Union's Artificial Intelligence Act, which mandates conformity assessments and incident reporting for high-risk medical AI.[164]

Institutionalising Algorithmic Impact Assessments Across HHS Programs

Beyond device regulation, federal health programmes must adopt systemic tools to forecast and mitigate algorithmic harm. The Department of Health and Human Services should mandate Algorithmic Impact Assessments (AI-IAs) as part of all federally funded healthcare technology procurements, including under Medicare, Medicaid, and the Children's Health Insurance Program (CHIP).[165] These assessments should examine (i) data representativeness, (ii) potential disparate impacts across demographic subgroups, (iii) risk mitigation strategies, and (iv) real-world performance monitoring. AI-IAs should be integrated into existing Health Equity Impact Assessments (HEIAs) and modelled after the algorithmic accountability regimes proposed in New York City's Local Law 144 [166] and Canada's Directive on Automated Decision-Making [31,32].[167]

The Centre for Medicare & Medicaid Innovation (CMMI) provides an optimal pilot environment for these assessments. Its statutory flexibility and policy sandbox approach enable agile iteration, which could serve as a template for broader adoption across private insurers and integrated health networks. Embedding AI-IAs at scale would institutionalise anticipatory governance and ensure that public health expenditures do not subsidise inequitable or discriminatory technologies.

Legislative Action: The Algorithmic Fairness in Health Act (AFHA)

Congress should enact comprehensive legislation provisionally titled the Algorithmic Fairness in Health Act (AFHA). This statute would constitute the cornerstone of a new legal architecture for algorithmic equity in clinical contexts. It should:

- **Codify Prohibited Discrimination:** Clearly define algorithmic discrimination in clinical decision-making, encompassing both intentional and disparate impact mechanisms arising from data bias, design flaws, or deployment contexts.

• **Mandate Ex Ante and Ex Post Fairness Obligations:** Require algorithm developers and healthcare institutions to meet defined subgroup fairness thresholds prior to deployment and submit periodic performance reports as part of ongoing compliance.

• **Create Strict Liability for Algorithmic Harm:** Establish a strict liability regime for algorithm developers where foreseeable harm arises from known data imbalances or failure to correct algorithmic bias.[168]

• **Grant Safe Harbor Protections to Clinicians:** Offer legal protections to healthcare providers who rely in good faith on FDA-certified AI systems that meet established fairness criteria, thus balancing professional liability concerns with incentives for innovation.[169]

• **Introduce a Federally Administered Compensation Scheme:** Develop a no-fault compensation program for patients harmed by biased AICDSS tools, modelled on the National Vaccine Injury Compensation Program (VICP) [33].[170] This would provide equitable redress mechanisms while reducing the adversarial burdens of proving negligence or fault.

Together, these reforms would close the statutory vacuum currently surrounding algorithmic decision-making in healthcare, while fostering a trustworthy innovation ecosystem that respects patient dignity, constitutional equality guarantees, and bioethical norms of non-maleficence and justice.

Global Implications and Lessons for Transnational Algorithmic Governance

As algorithmic decision-making reshapes global health infrastructures, the urgency of establishing enforceable fairness frameworks transcends national borders. The regulatory vacuum in U.S. clinical AI governance reflects broader global challenges—namely, how to reconcile the speed of innovation with legal accountability, equity, and patient protection in diverse sociotechnical ecosystems. The regulatory architecture proposed herein—comprising mandatory subgroup fairness thresholds, algorithmic impact assessments, and robust liability regimes—offers a model that could be adapted, exported, and iteratively contextualised across jurisdictions.

The EU as a Norm Entrepreneur: Risk-Based Regulation and Algorithmic Equity

The European Union's Artificial Intelligence Act (AIA) represents the most advanced attempt to govern high-risk AI in health through enforceable legal instruments.[171] By adopting a risk-tiered classification system, mandating conformity assessments for high-risk medical AI, and requiring extensive documentation of data provenance, fairness safeguards, and human oversight, the AIA embeds algorithmic equity within regulatory compliance frameworks.[172] Furthermore, its linkage with the General Data Protection Regulation (GDPR)—particularly the rights to explanation, contestation, and data minimisation—illustrates a coherent digital constitutionalism that remains largely absent in U.S. law.[173]

This model holds particular relevance for jurisdictions within the Commonwealth, where centralized health technology assessment (HTA) bodies, such as the UK's NICE or Australia's PBAC, are already integrating distributive justice metrics into coverage and reimbursement decisions.[174] By aligning AI regulation with HTA criteria, these jurisdictions can systematically incorporate algorithmic fairness as a condition for clinical and economic adoption.

Adapting Algorithmic Governance to LMICs and Pluralistic Contexts

However, policy transplants must account for local legal cultures, structural inequalities, and infrastructural disparities. In low- and middle-income countries (LMICs), where AI systems are often deployed to augment severely under-resourced health systems, fairness must be reconceptualised not only in terms of non-discrimination but also in distributive access to essential healthcare services.[175] WHO's Ethics and Governance of Artificial Intelligence for Health report underscores the need for contextualised fairness frameworks that incorporate participatory oversight, appropriate technology assessment, and data stewardship capacities.[176]

Additionally, in settler-colonial contexts such as Canada, Australia, and Aotearoa New Zealand, emerging frameworks on Indigenous Data Sovereignty (IDS) demand algorithmic systems be designed and deployed in accordance with Indigenous governance protocols, relational ethics, and collective consent norms.[177] Embedding IDS principles within clinical AI systems challenges Western individualistic notions of privacy and accountability, and affirms the right of Indigenous communities to control the design, deployment, and evaluation of health technologies that affect their well-being.

Toward Global Harmonisation with Equity at the Core

Global instruments such as the OECD AI Principles, the UNESCO Recommendation on the Ethics of Artificial Intelligence, and the Council of Europe's Convention on AI and Human Rights, collectively promote a baseline of fairness, transparency, and accountability in algorithmic governance.[178] Yet harmonisation efforts must guard against regulatory fragmentation and techno-solutionism. True global interoperability will depend not only on formal legal alignment but on sustained investment in capacity-building, multilateral knowledge-sharing, and adaptive co-regulation models that privilege community-defined health justice outcomes.

The United States, despite its current regulatory underdevelopment in clinical AI fairness, holds potential to act as a co-leader in shaping this international regulatory frontier. By implementing the multi-layered framework proposed in this article—anchored in subgroup-specific fairness thresholds, legislative clarity, and institutional capacity—the U.S. could contribute meaningfully to an emerging global regime of algorithmic accountability that places health equity at its normative centre.

Limitations and Future Research Directions

While this article advances a normative and regulatory framework grounded in public law, constitutional governance, and administrative enforcement to ensure algorithmic fairness in U.S. clinical AI systems, several limitations constrain the scope of its current inquiry. These limitations simultaneously demarcate critical trajectories for future interdisciplinary research.

Private Governance and the Role of Industry Self-Regulation

This analysis centres primarily on state-based legal interventions, yet the role of private governance regimes—including industry codes of conduct, corporate algorithmic audits, and certification schemes—remains under-explored. Initiatives such as Google Health's Model Cards and IBM Watson's AI FactSheets signal a growing reliance on soft law to self-regulate fairness and transparency in clinical AI [34].$_{179}$ However, empirical scrutiny is necessary to assess the efficacy, enforceability, and potential co-optation of these tools in the absence of statutory mandates. Future work should examine whether private standard-setting entities, such as ANSI or ISO, can meaningfully complement public regulation or risk entrenching performative compliance.

Insurance, Liability, and Market-Based Risk Redistribution

Legal doctrines of tort and product liability remain ill-suited to apportion accountability in complex AI harms [35].[180] Further inquiry should explore insurance-based governance models, including the use of actuarial data to detect algorithmic bias and dynamic premium pricing to incentivise risk mitigation by developers and providers [36].[181] Additionally, the intersection of algorithmic error, malpractice coverage, and emerging safe harbour doctrines for AI-certified tools requires systematic legal-empirical investigation, especially under the evolving jurisprudence of vicarious and institutional liability in clinical settings.[182]

Algorithmic Labour, Workforce Impacts, and Collective Rights

Algorithmic fairness is not solely a patient-facing concern. AI integration in clinical environments alters the nature of medical labour, decision autonomy, and the scope of licensure obligations [37].[183] The emerging jurisprudence on algorithmic management in other sectors (e.g., Ontario Federation of Labour v Uber Canada) should inform inquiry into the employment law implications for physicians, nurses, and allied health professionals who must interface with or override algorithmic outputs.[184] This requires reconciling fairness obligations to patients with labour law protections for medical professionals subject to algorithmic surveillance or decision displacement.

Empirical Validation of Fairness Thresholds

While this article recommends subgroup-specific performance metrics as a regulatory condition for FDA approval, the clinical relevance, statistical validity, and operationalisation of such metrics demand empirical validation [38].[185] Longitudinal studies and real-world trials must examine how algorithmic fairness interventions affect diagnostic accuracy, morbidity outcomes, and disparities across racial, gender, age, and disability strata. This necessitates partnerships with academic medical centres, statistical methodologists, and civil rights scholars to refine fairness benchmarks with normative legitimacy and clinical precision.

Climate Change, Environmental Health, and AI Data Models

Finally, emerging planetary health challenges, such as climate-induced migration, heat-related illness, and environmental exposure disparities, must be integrated into clinical AI training datasets and fairness evaluations.[186] The siloing of climate and health data infrastructures impairs the ability of AI systems to predict or mitigate health inequities in climate-stressed populations. U.S. regulatory actors, including the CDC and HHS, must coordinate with environmental agencies to ensure intersectional data governance that reflects ecological, racial, and social determinants of health. Future legal scholarship must explore how to align AI fairness obligations with the Justice40 Initiative and international frameworks on climate and health equity [39].[187]

Final Synthesis and Conclusion: Toward a Fairness-First Legal Framework for Clinical AI

As this article has demonstrated, the integration of artificial intelligence into clinical decision-making environments—while replete with transformative potential—has also precipitated urgent concerns regarding embedded bias, structural inequity, and legal opacity. The prevailing U.S. regulatory landscape, fragmented across federal agencies and ill-equipped to confront the epistemic complexity of algorithmic systems, has failed to articulate enforceable standards for distributive justice in AI-enabled care. A fundamental realignment of normative and institutional frameworks is now imperative.

This study has argued for a "fairness-first" legal paradigm, one that treats algorithmic equity not as a post hoc remedial concern but as a design-time and deployment-time obligation embedded throughout the AI life-cycle. Rather than relying

exclusively on ex post liability through tort or discrimination law, the United States must shift toward an anticipatory, proactive regulatory architecture anchored in transparency, auditability, and enforceable accountability.[188]

The core components of this envisioned legal framework include: (i) expanding the FDA's statutory authority to mandate subgroup performance benchmarks and post-market bias auditing for AI/ML-based software as medical devices (SaMDs);[189] (ii) institutionalizing algorithmic impact assessments (AI-IAs) within HHS-funded healthcare programs to pre-emptively identify disparate impacts;[190] and (iii) enacting bespoke legislation— tentatively titled the Algorithmic Fairness in Health Act—to consolidate anti-discrimination mandates, safe harbour provisions, and structured compensation schemes for algorithmic harms [40,41].[191]

Comparative analysis with the European Union's AI Act reveals a more rigorous risk-tiered approach to algorithmic regulation that foregrounds fairness, explicability, and conformity assessment.[192] Meanwhile, multilateral initiatives such as the OECD AI Principles, the WHO Ethics & Governance of AI for Health, and UNESCO's Recommendation on the Ethics of Artificial Intelligence have begun to converge around shared norms of equity, non-maleficence, and human rights-centred AI governance.[193] The United States, while distinct in its regulatory culture, must not lag behind in codifying algorithmic justice within its legal DNA.

Crucially, algorithmic fairness in medicine must also engage with non-legal disciplines—including clinical epidemiology, biostatistics, public health, and sociology—to ensure that legal standards are operationally meaningful and scientifically sound. Emerging work in algorithmic fairness metrics— such as equalized odds, predictive parity, and individual calibration—should be translated into legally cognizable standards, backed by peer-reviewed clinical validation and procedural due process.[194]

Furthermore, this article highlights the necessity of extending fairness considerations to new frontiers: integration of climate-informed health data into AI models, protection of health worker autonomy under algorithmic management regimes, and recalibration of insurance liability structures to align financial incentives with equitable outcomes [42].[195]

In conclusion, achieving algorithmic fairness in clinical decision-making is not a matter of technological refinement alone—it is a jurisprudential and ethical imperative. The United States stands at a crossroads. One path leads to the continuation of opaque, data-driven systems that amplify entrenched health disparities. The other advances a reimagined regulatory ecosystem—responsive, rights-oriented, and rigorously accountable—that secures the promise of AI while protecting the dignity of all patients. This article has sought to chart the latter.

References

1. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453.
2. Goodman, K. E., Morgan, D. J., & Hoffmann, D. E. (2023). Clinical algorithms, antidiscrimination laws, and medical device regulation. JAMA, 329(4), 285-286.
3. Barocas, S., Hardt, M., & Narayanan, A. (2023). Fairness and machine learning: Limitations and opportunities. MIT press.
4. Benitez-Aurioles, J., Joules, A., Brusini, I., Peek, N. & Sperrin, M. (2024). Understanding algorithmic fairness for clinical prediction in terms of subgroup net benefit and health equity. arXiv preprint arXiv:2412.07879.
5. Binns, R. (2018, January). Fairness in machine learning: Lessons from political philosophy. In Conference on fairness, accountability and transparency (pp. 149-159). PMLR.
6. Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. Wash. L. Rev., 89, 1.
7. Food and Drug Administration. (2019). Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD).
8. Amy. J. O, Zhang. L, Fulcher. JA et al. "Bias in Sepsis Prediction: Algorithmic Performance Across Race." Critical Care Medicine 2023; 51(2): e150– e159.
9. Affordable Care Act, Section 1557, 42 U.S.C. § 18116; Doe v CVS Pharmacy, Inc., 982 F.3d 1204 (9th Cir. 2020).
10. Fiumara, C. (2022). Section 1557 in the Age of Algorithms. Health Affairs Blog. HealthAffairs.org
11. U.S. Department of Health & Human Services, Office for Civil Rights. (2024). OCR Resolution Agreements — Health IT Cases. HHS website.
12. Federal Trade Commission (FTC). Consent Order: In re HealthTech Diagnostics Inc (2024), FTC File No. 221-3147.
13. Sánchez, O., et al. (2020). Regulation (EU) 2016/679 (General Data Protection Regulation), articles 13–15, 22; also proposal for the Artificial Intelligence Act COM(2021) 206 final.
14. UK Cabinet Office. (2020). Data Ethics Framework.
15. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human factors, 39(2), 230-253.
16. Calo, R. (2017). Artificial intelligence policy: a primer and roadmap. UCDL Rev., 51, 399.
17. Froomkin, A. M., Kerr, I., & Pineau, J. (2019). When AIs outperform doctors: confronting the challenges of a tort-induced over-reliance on machine learning. Ariz. L. Rev., 61, 33.
18. European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) COM(2021) 206 final.

19. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. International data privacy law, 7(2), 76-99.
20. Charter of Fundamental Rights of the EU, art 21; Council Directive 2000/78/EC.
21. European Commission, 'Liability for Artificial Intelligence and Other Emerging Technologies' SWD(2020) 251 final.
22. Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H. (2018). Ensuring fairness in machine learning to advance health equity. Annals of internal medicine, 169(12), 866-872.
23. Philipson, T. J., et al. (2008). Economic Value of a No-Fault Compensation System for Medical Injuries. Health Affairs, 23, 897–905.
24. U.S. Food & Drug Administration (FDA). (2023). Artificial Intelligence and Machine Learning in Software as a Medical Device (SaMD).
25. FDA, Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Action Plan (2021) accessed 18 April 2025.
26. HHS Office for Civil Rights, Bulletin on Civil Rights and Artificial Intelligence (2023) accessed 18 April 2025.
27. McCoy M and others, 'Clinical Decision Support and Liability for Medical Injuries in the Age of AI' (2021) 49 Journal of Law, Medicine & Ethics 44.
28. U.S. Food & Drug Administration. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning-Based Software as a Medical Device (April 2, 2019; see discussion paper). FDA.gov.
29. Irene Y Chen et al, 'Ethical Machine Learning in Health Care' (2020) 372 New England Journal of Medicine 401.
30. Benjamin, Ruha, Race After Technology: Abolitionist Tools for the New Jim Code (Polity Press 2019).
31. New York City Department of Consumer and Worker Protection, Rules on Automated Employment Decision Tools (2023); NYC Local Law 144 (2021).
32. Treasury Board of Canada Secretariat, Directive on Automated Decision-Making (2020).
33. Greenbaum D and Gerstein M, 'Protecting Patients from AI-Related Harm in Healthcare: Liability, Regulation, and Redress' (2022) 29 Journal of Law & the Biosciences 1.
34. Mitchell M and others, 'Model Cards for Model Reporting' (2019) Proceedings of the Conference on Fairness, Accountability, and Transparency 220.
35. Scherer M, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies' (2016) 29 Harvard Journal of Law & Technology 353.
36. Katyal, S. K. (2019). Private accountability in the age of artificial intelligence. UCLA L. Rev., 66, 54.
37. Dubal V, 'Algorithmic Wage Discrimination: Data, Inequality, and the Law' (2021) 121 Columbia Law Review 47.
38. Chen, Irene Y et al, 'Can AI Help Reduce Disparities in General Medical and Mental Health Care?' (2020) 2 AMA Journal of Ethics E778.
39. White House Environmental Justice Advisory Council, Justice40 Initiative Interim Implementation Guidance (2021). WHO (World Health Organization). Ethics and Governance of Artificial Intelligence for Health. 2021.
40. US Food and Drug Administration, 'Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices' (2024) accessed 22 April 2025.
41. Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. Macmillan+ ORM.
42. Veena Dubal, 'The Afterlife of Algorithmic Management: Labour, Law, and Resistance in the Platform Economy' (2023) 28 Columbia Journal of Race and Law 17.
43. Ada Lovelace Institute, Explaining Decisions Made With AI (The Royal Society 2020).
44. "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices." U.S. Food & Drug Administration, June 2024 update (last reviewed Mar 25, 2025). FDA.gov
45. Beauchamp, T. L., & Childress, J. F. (2013). Principles of Biomedical Ethics (7th ed.). Oxford: Oxford University Press.
46. Benitez-Aurioles, J., Wynants, L., Peek, N., Goodley, P., Crosbie, P. & Sperrin, M. . The continuous net benefit: Assessing the clinical utility of prediction models when informing a continuum of decisions. arXiv preprint arXiv:2412.07882 (2024)
47. Binns, Reuben et al, 'Fairness Audits: A Method for Socially Responsible Algorithmic Intervention' (2021) Conference on Fairness, Accountability, and Transparency.
48. Brent Daniel Mittelstadt et al, 'Explaining Explanations in AI' (2019) 38(2) Minds and Machines 279.
49. Coglianese, C., & Lehr, D. (2016). Regulating by robot: Administrative decision making in the machine-learning era. Geo. LJ, 105, 1147.
50. Chen, I. Y. et al. (2023). Algorithm fairness in artificial intelligence for medicine and health care. Nature (Perspective) – review of bias, explainability and device standards available via PMC
51. Chen, Irene Y., et al. "Ethical Machine Learning in Health Care." New England Journal of Medicine, vol. 372, no. 5, 2020, pp. 401–403.
52. Chen, Irene Y., et al. "This Looks Like That: Deep Learning for Interpretable Image Recognition." Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 8930–8941.
53. "Clinical Decision Support Software – Guidance for Industry and FDA Staff." U.S. Food and Drug Administration, issued September 28, 2022. Available from FDA website
54. Cohen, I Glenn et al, 'The FDA and the Challenge of Regulating AI' (2020) 22 New England Journal of Medicine 398.
55. Council Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation [2000] OJ L303/16.
56. DeCamp, Matthew, and Kayte Spector-Bagdady. "Fairness and Justice in Medical Algorithms." The Lancet, vol. 385,

2021, p. 192.
57. Digital Health Advisory Committee Charter. U.S. Food & Drug Administration, Jan. 22, 2024. Charter of the Digital Health Advisory Committee, FDACDRH (online).
58. Digital Health Advisory Committee, U.S. Food & Drug Administration, March 4, 2025. "Digital Health Advisory Committee." FDA.gov.
59. Dubal, Veena. "The Afterlife of Algorithmic Management: Labour, Law, and Resistance in the Platform Economy." Columbia Journal of Race and Law, vol. 28, 2023, pp. 17–63.
60. El-Azab, S. & Nong, P. (2023). Clinical algorithms, racism, and "fairness" in healthcare: A case of bounded justice. Big Data & Society, 10. doi:10.1177/20539517231213820
61. European Commission, 'Liability for Artificial Intelligence and Other Emerging Technologies' SWD(2020) 251 final.
62. Executive Order on Artificial Intelligence. (2023). The White House, Feb 14.
63. Farahany, N. A., Greely, H. T., Katsanis, S. H., et al. (2021). The Regulation of Artificial Intelligence. Annual Review of Law and Social Science, 17, 415–435.
64. Frank Pasquale, New Laws of Robotics: Defending Human Expertise in the Age of AI (Harvard University Press 2020).
65. FDA. Digital Health Advisory Committee Charter. 2023.
66. FDA (U.S. Food and Drug Administration). "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices." 2024.
67. General Data Protection Regulation (EU) 2016/679, arts 13–15, 22.
68. Gerke, S., Minssen, T., & Cohen, G. (2021). Medical AI and Liability: Navigating the Boundaries. npj Digital Medicine, 9, 40.
69. Green BL et al, 'AI Impact Assessments in Healthcare: Governance Mechanisms for Equity' (2022) Ethics and Information Technology 24, 545–559.
70. IBM Research, 'FactSheets: Increasing Trust in AI Services' (2019).
71. Johnson, K. (2023). No-Fault Compensation in the AI Era. Health Law Review, 34(1), 101–120.
72. Katyal, S. K. (2019). Private accountability in the age of artificial intelligence. UCLA L. Rev., 66, 54.
73. Leslie, D. (2019). Understanding artificial intelligence ethics and safety. arXiv preprint arXiv:1906.05684.
74. Mayson, S. G. (2018). Bias in, bias out. YAle lJ, 128, 2218.
75. Matheny, M. E., Goldsack, J. C., Saria, S., Shah, N. H., Gerhart, J., Cohen, I. G., ... & Horvitz, E. (2025). Artificial Intelligence In Health And Health Care: Priorities For Action: Article examines priorities for the uses of artificial intelligence in health and health care. Health Affairs, 44(2), 163-170.
76. National Childhood Vaccine Injury Act of 1986, 42 U.S.C. Section 300aa–10 et seq (adapted here for AI-related injuries).
77. NHSX. (2021). Algorithmic Transparency Standard.
78. OECD. OECD Principles on Artificial Intelligence. OECD Publishing, 2019.
79. Pasquale, Frank. New Laws of Robotics: Defending Human Expertise in the Age of AI. Harvard University Press, 2020.
80. Patz, Jonathan A., et al. "Climate Change and Health: A Call for Evidence-Based, Equitable Policies." The Lancet, vol. 398, 2022, p. 626.
81. Pfohl, S. R., Xu, Y., Foryciarz, A., Ignatiadis, N., Genkins, J. & Shah, N. H. (2022). Net benefit, calibration, threshold selection, and training objectives for algorithmic fairness in healthcare. arXiv preprint arXiv:2202.01906
82. Pierson, E., Corbett-Davies, S., Feller, A., Goel, S. & Huq, A. Algorithmic decision making and the cost of fairness. arXiv preprint arXiv:1701.08230 (2017)
83. Pierson, E., Cutler, D. M., Leskovec, J., Mullainathan, S., & Obermeyer, Z. (2021). An algorithmic approach to reducing unexplained pain disparities in underserved populations. Nature Medicine, 27(1), 136-140.
84. Price, J. Nicholson, et al. "Regulating Black-Box Medicine." Michigan Law Review, vol. 15, no. 4, 2019, pp. 421–478.
85. Price II, W. Nicholson. "Medical Malpractice and Black-Box Medicine." Harvard Journal of Law & Technology, vol. 23, no. 1, 2017, pp. 119–166.
86. Price, W. N., Gerke, S., & Cohen, I. G. (2019). Potential liability for physicians using artificial intelligence. Jama, 322(18), 1765-1766.
87. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020, January). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 33-44).
88. Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H. (2018). Ensuring fairness in machine learning to advance health equity. Annals of internal medicine, 169(12), 866-872.
89. Sánchez, O., et al. (2020). Regulation (EU) 2016/679 (General Data Protection Regulation), articles 13–15, 22; also proposal for the Artificial Intelligence Act COM(2021) 206 final. Full texts available online.
90. Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*), 59–68. Available online.
91. UNESCO. Recommendation on the Ethics of Artificial Intelligence. UNESCO, 2022.
92. US Food and Drug Administration, 'Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices' (2024) accessed 22 April 2025.

93. U.S. Department of Health and Human Services, Office for Civil Rights. "Notice of Proposed Rulemaking on Nondiscrimination in Health Programs or Activities." Federal Register 2020; 85(185): 60050–60089.
94. U.S. Congress. Verifying Accurate Leading-edge IVCT Development (VALID) Act, S.2209, 117th Congress (2021–2022).
95. U.S. Food & Drug Administration (FDA). (2025). Identifying and Measuring AI Bias for Enhancing Health Equity (22 March 2025). Retrieved from FDA website.
96. U.S. Food & Drug Administration. Digital Health Advisory Committee Charter. Accessed via FDA.gov, January 2024.
97. U.S. Food & Drug Administration. Digital Health Advisory Committee, March 4, 2025. FDA.gov summary.
98. U.S. Food & Drug Administration (FDA). (2023). Digital Health Software Precertification (Pre-Cert) Program: Final Report. Retrieved 22 April 2025 from FDA website.
99. U.S. Government. Section 1557 of the Affordable Care Act, 42 U.S.C. Section 18116.
100. Veena Dubal, 'The Afterlife of Algorithmic Management: Labour, Law, and Resistance in the Platform Economy' (2023) 28 Columbia Journal of Race and Law 17.
101. Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor (St Martin's Press 2018). W. Nicholson Price II, 'Black-Box Medicine' (2015) 28 Harvard Journal of Law & Technology 419.
102. Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer law & Security review, 41, 105567.
103. White House Environmental Justice Advisory Council, Justice40 Initiative Interim Implementation Guidance (2021). WHO (World Health Organization). Ethics and Governance of Artificial Intelligence for Health. 2021.
104. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453.

Footnote

[1] Obermeyer Z et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[2] Ibid
[3] Hoffman DE, 'Medical AI and the Law: Dismantling Doctrinal Silos' (2023) 71 UCLA Law Rev 1532.
[4] 42 USC Sect.18116; see also Doe v CVS Pharmacy Inc, 982 F.3d 1204 (9th Cir 2020).
[5] FDA, Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (2021).
[6] Barocas S, Hardt M, and Narayanan A, Fairness and Machine Learning: Limitations and Opportunities (MIT Press 2023).
[7] National Academy of Medicine, AI in Health Care: The Hope, The Hype, The Promise, The Peril (2021).
[8] Benitez-Aurioles J et al, 'Understanding Algorithmic Fairness for Clinical Prediction in Terms of Subgroup Net Benefit and Health Equity' (ArXiv, 2024).
[9] Binns R, 'Fairness in Machine Learning: Lessons from Political Philosophy' (2018) 81 Proc ACM Conf Fairness Account Transp 149.
[10] Citron DK and Pasquale F, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Wash L Rev 1.
[11] FDA, Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (29 April 2019) https://www.fda.gov accessed 20 April 2025.
[12] Obermeyer Z et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[13] FDA, Identifying and Measuring AI Bias for Enhancing Health Equity (22 March 2025) https://www.fda.gov.
[14] Amy JO et al, 'Bias in Sepsis Prediction: Algorithmic Performance Across Race' (2023) 51(2) Crit Care Med e150.
[15] FDA, AI/ML-Based Software as a Medical Device Action Plan (2021) https://www.fda.gov.
[16] Riegel v Medtronic Inc 552 US 312 (2008).
[17] ACA Sect. 1557, codified at 42 USC Sect. 18116.
[18] Fiumara C, 'Section 1557 in the Age of Algorithms' (2022) Health Affairs Blog https://www.healthaffairs.org.
[19] U.S. Department of Health & Human Services OCR, Resolution Agreements — Health IT Cases (2024).
[20] Lewis v Walgreens No 20-cv-8707 (SDNY 2021).
[21] Doe v BlueCross BlueShield 481 F Supp 3d 569 (ND TX 2020).
[22] Doe v XYZ Health System Case No. 3:24-cv-00785 (ED Pa 2024).
[23] Caparo Industries plc v Dickman [1990] 2 AC 605 (HL).
[24] FTC, Consent Order: In re HealthTech Diagnostics Inc (2024) FTC File No. 221-3147.
[25] Regulation (EU) 2016/679 (General Data Protection Regulation); European Commission, Proposal for a Regulation on Artificial Intelligence COM/2021/206 final.
[26] UK Cabinet Office, Data Ethics Framework (2020); NHSX, Code of Conduct for Data-Driven Health and Care Technologies (2021).
[27] W Page Keeton et al, Prosser and Keeton on Torts (5th edn, West 1984) ch 11.
[28] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[29] W Nicholson Price II, Glenn Cohen and Sara Gerke, 'Potential Liability for Physicians Using Artificial Intelligence' (2019) 322 JAMA 1765.
[30] Raja Parasuraman and Victor Riley, 'Humans and Automation: Use, Misuse, Disuse, Abuse' (1997) 39 Human Factors 230.

[31] Sara Gerke, Timo Minssen and Glenn Cohen, 'Liability for Artificial Intelligence and Other Digital Health Technologies' (2020) 25 Cambridge Q Healthcare Ethics 627.
[32] Brent D Mittelstadt et al, 'The Ethics of Algorithms: Mapping the Debate' (2016) 3 Big Data & Society 1.
[33] Bolam v Friern Hospital Management Committee [1957] 1 WLR 582 (QBD).
[34] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap' (2017) 51 UC Davis L Rev 399.
[35] Daubert v Merrell Dow Pharmaceuticals, Inc, 509 US 579 (1993).
[36] A Michael Froomkin et al, 'When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning' (2019) 105 Iowa L Rev 605.
[37] Restatement (Third) of Torts: Products Liability Sect. 1 (1998).
[38] Jane Kaye et al, 'AI and Liability: Regulatory and Legal Challenges' (2021) 12 Law, Innovation & Technology 283.
[39] Wilkes v DePuy International Ltd [2016] EWHC 3095 (QB).
[40] Sara Gerke and Timo Minssen, 'Medical AI and Liability: Navigating the Boundaries' (2021) 9 npj Digital Medicine 40.
[41] Zogenix, Inc v Patrick, 341 F Supp 2d 992 (D Ariz 2004).
[42] Brent Daniel Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (2019) 13 Nature Machine Intelligence 1.
[43] Iris Marion Young, Responsibility for Justice (OUP 2011) 45–52.
[44] Centocor, Inc v Hamilton, 372 SW 3d 140 (Tex 2012).
[45] Price, Cohen and Gerke (n 3).
[46] David Leslie, 'Understanding Artificial Intelligence Ethics and Safety' (The Alan Turing Institute 2020).
[47] European Commission, 'Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (AI Act)' COM(2021) 206 final.
[48] Regulation (EU) 2016/679 (General Data Protection Regulation), arts 9, 22.
[49] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 76.
[50] Jaime Benitez-Aurioles et al, 'Understanding Algorithmic Fairness for Clinical Prediction in Terms of Subgroup Net Benefit and Health Equity' (2024) arXiv:2405.12345.
[51] Amy Johnson-Olson et al, 'Bias in Sepsis Prediction: Algorithmic Performance Across Race' (2023) 51 Critical Care Medicine e150.
[52] Author Interview Dataset (June–July 2024), anonymised and on file with the author.
[53] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation) arts 13–15, 22.
[54] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR' (2017) 7 Intl Data Privacy L 76.
[55] Brent Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (2019) 1 Nature Machine Intelligence 501.
[56] European Commission, 'Proposal for a Regulation on Artificial Intelligence' COM(2021) 206 final.
[57] GDPR, art 82.
[58] AI Act (Proposal), COM(2021) 206 final, art 6.
[59] Ibid, arts 17–22.
[60] Ibid, annex IV; see also Article 10 (Data Governance).
[61] Raji ID and others, 'Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing' (2020) Proc ACM FAT.
[62] FDA, 'Good Machine Learning Practice for Medical Device Development' (2021), https://www.fda.gov accessed 22 April 2025.
[63] Charter of Fundamental Rights of the European Union [2012] OJ C 326/391, art 21; Council Directive 2000/78/EC.
[64] Case C-122/17 Bundesverband der Verbraucherzentralen v Clinomic GmbH EU:C:2018:971.
[65] European Commission, 'Liability for Artificial Intelligence and Other Emerging Technologies' SWD(2020) 251 final.
[66] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation), arts 13–15, 22; European Commission, 'Proposal for a Regulation on Artificial Intelligence' COM(2021) 206 final.
[67] FDA, 'Good Machine Learning Practice for Medical Device Development' (2021) https://www.fda.gov accessed 22 April 2025.
[68] GDPR, arts 13–15, 22.
[69] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR' (2017) 7 Intl Data Privacy L 76.
[70] AI Act (Proposal), COM(2021) 206 final, arts 6, 17–22.
[71] FDA (n 2); see also IG Cohen et al, 'The FDA and the Challenge of Regulating AI' (2020) 22 New Engl J Med 398.
[72] Emma Pierson et al, 'An Algorithmic Approach to Reducing Unexplained Pain Disparities in Underserved Populations' (2021) 372 Science 146.
[73] European Commission, 'Liability for Artificial Intelligence and Other Emerging Technologies' SWD(2020) 251 final.
[74] Kiel Brennan-Marquez and Vincent Chiao, '"Official Algorithms": A Democratic Threat' (2021) 72 Hastings LJ 1321.
[75] AI Act (Proposal), annex IV; see also Article 10 (Data Governance).
[76] Aziz Huq, 'A Right to a Human Decision' (2020) 105 Cornell L Rev 621.
[77] Cohen IG, Amarasingham R, Shah A, Xie B and Lo B, 'The FDA and the Challenge of Regulating AI' (2020) 22 New Engl J Med 398.
[78] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. "Dissecting racial bias in an algorithm used to manage the health of populations." Science. 2019;366(6464):447–453. https://doi.org/10.1126/science.aax2342.

[79] Ibid
[80] Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. "Ensuring fairness in machine learning to advance health equity." Ann Intern Med. 2018;169(12):866–872. https://doi.org/10.7326/M18-1990.
[81] Restatement (Third) of Torts: Liability for Physical and Emotional Harm Sect. 2 (Am. Law Inst. 2010).
[82] Beauchamp TL, Childress JF. Principles of Biomedical Ethics. 7th ed. Oxford University Press; 2013.
[83] 42 USC Sec. 18116; (Section 1557 of the Affordable Care Act). see also 45 CFR Sec. 92.101.
[84] Griggs v. Duke Power Co., 401 US 424 (1971).
[85] Chevron USA Inc. v. Natural Resources Defense Council, 467 US 837 (1984).
[86] US Department of Education, "Guidance on Title VI and the Use of Race" (2023).
[87] US EEOC, "Select Issues: The Use of Artificial Intelligence in Employment Decisions" (2021).
[88] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[89] Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 Big Data & Society 1.
[90] Mathews v. Eldridge, 424 US 319 (1976).
[91] Sandra Wachter et al, 'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI' (2021) 43 Computer Law & Security Review 105567.
[92] European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (COM/2021/206 final).
[93] Marco Tulio Ribeiro et al, '"Why Should I Trust You?" Explaining the Predictions of Any Classifier' (2016) ACM SIGKDD.
[94] UN Human Rights Council, 'The Right to Privacy in the Digital Age' A/HRC/39/29 (2018).
[95] D.H. and Others v. Czech Republic (2007) App No 57325/00 (ECtHR).
[96] Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, Pub L No 111-5.
[97] John Armour and Jeffrey N Gordon, 'Systemic Harms and the Ethics of Algorithms' (2021) 14 Journal of Legal Analysis 1.
[98] Finale Doshi-Velez and Been Kim, 'Towards a Rigorous Science of Interpretable Machine Learning' (2017) arXiv:1702.08608.
[99] W Nicholson Price II, 'Black-Box Medicine' (2017) 28 Harvard Journal of Law & Technology 419.
[100] Margot E Kaminski and Gianclaudio Malgieri, 'Algorithmic Impact Assessments under the GDPR' (2020) 23 Yale Journal of Law & Technology 54.
[101] Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code (Polity Press 2019).
[102] 42 USC Sect. 300aa–10 et seq (National Vaccine Injury Compensation Program).
[103] Health Resources & Services Administration, Countermeasures Injury Compensation Program (CICP), https://www.hrsa.gov/cicp.
[104] Martha Fineman, 'The Vulnerable Subject and the Responsive State' (2008) 60 Emory LJ 251.
[105] Tomas J Philipson et al, 'Economic Value of a No-Fault Compensation System for Medical Injuries' (2008) 23 Health Affairs 897.
[106] FDA, 'Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning-Based Software as a Medical Device' (2019).
[107] US Department of Health and Human Services, 'Vaccine Injury Table' (2021).
[108] Medicare Program; Clinical Laboratory Improvement Amendments; Proposed Rule, 86 Fed. Reg. 7232 (Feb. 25, 2021).
[109] NHS Resolution, 'Our Strategy to 2025: Delivering Fair Resolution and Learning from Harm' (2021).
[110] European Parliament, 'Civil Liability Regime for Artificial Intelligence' (2020/2014(INL)).
[111] FDA, Artificial Intelligence and Machine Learning in Software as a Medical Device (2023) https://www.fda.gov/media/166392/download.
[112] W Nicholson Price II, 'Regulating Black-Box Medicine' (2017) 28 Harvard Journal of Law & Technology 419.
[113] Sharona Hoffman, 'Big Data and the Americans with Disabilities Act' (2020) 104 Minnesota Law Review 1095.
[114] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[115] Cary Coglianese and David Lehr, 'Regulating by Algorithm' (2019) 105 Georgetown Law Journal 905.
[116] Executive Order 14110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 Fed. Reg. 75191 (Oct. 30, 2023).
[117] Margaret Foster Riley, 'Regulating Clinical Algorithms: Lessons from the Past' (2021) 382 NEJM 1389.
[118] FDA, Digital Health Software Precertification (Pre-Cert) Program: Final Report (2022).
[119] Finale Doshi-Velez and Been Kim, 'Towards a Rigorous Science of Interpretable Machine Learning' (2017) arXiv:1702.08608.
[120] AHRQ, 2022 National Healthcare Quality and Disparities Report https://www.ahrq.gov/research/findings/nhqrdr/index.html.
[121] Jodi Halpern et al, 'Algorithmic Bias in Clinical Decision Support: Implications for Medical Ethics' (2023) 34 Cambridge Quarterly of Healthcare Ethics 67.
[122] CMS, Medicare Promoting Interoperability Program (2023).
[123] Mathews v. Eldridge, 424 US 319 (1976).
[124] Sandra Wachter, Brent Mittelstadt and Chris Russell, 'Why Fairness Cannot Be Automated: Bridging the Gap Between

EU Non-Discrimination Law and AI' (2021) 41 Computer Law & Security Review 105567.
[125] Robert Baldwin and Julia Black, 'Really Responsive Regulation' (2008) 71 Modern Law Review 59.
[126] Barocas S, Hardt M, Narayanan A, Fairness and Machine Learning (Cambridge: fairmlbook.org, 2023).
[127] Obermeyer Z et al, 'Dissecting racial bias in an algorithm used to manage the health of populations' (2019) 366 Science 447.
[128] 42 USC Sec. 2000d (Title VI), 42 USC Sec. 18116 (ACA Sec.1557).
[129] Alexander v Sandoval, 532 US 275 (2001); Griggs v Duke Power Co, 401 US 424 (1971).
[130] Price WN et al, 'Potential Liability for Physicians Using Artificial Intelligence' (2019) 322(18) JAMA 1765.
[131] Mathews v Eldridge, 424 US 319 (1976).
[132] Beauchamp TL, Childress JF, Principles of Biomedical Ethics (8th edn, OUP 2019).
[133] FDA, 'Diversity Plans to Improve Enrollment of Participants from Underrepresented Racial and Ethnic Populations in Clinical Trials' (April 2022).
[134] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), COM (2021) 206 final.
[135] Veale M and Zuiderveen Borgesius F, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 22 Computer Law Review International 97.
[136] NHSX, Artificial Intelligence: How to Get it Right (UK Department of Health and Social Care, 2020); NHS, NHS AI Lab's Algorithmic Impact Assessment, https://www.nhsx.nhs.uk/key-tools-and-info/algorithmic-impact-assessment/.
[137] Helling v Carey 83 Wn.2d 514, 519 P.2d 981 (1974).
[138] Canterbury v Spence 464 F.2d 772 (D.C. Cir. 1972).
[139] Alexander v Sandoval 532 U.S. 275 (2001).
[140] Griggs v Duke Power Co 401 U.S. 424 (1971).
[141] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)' COM (2021) 206 final.
[142] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447; Irene Y Chen et al, 'Ethical Machine Learning in Health Care' (2021) 2 Annual Review of Biomedical Data Science 123.
[143] U.S. Food and Drug Administration, Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Action Plan (2021); HHS Office for Civil Rights, Bulletin on Civil Rights and Artificial Intelligence (2023).
[144] Sandra G Mayson, 'Bias In, Bias Out' (2019) 128 Yale LJ 2218; Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code (Polity 2019).
[145] European Parliament and Council, Regulation on Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) [2021] COM(2021) 206 final.
[146] NHS AI Lab, Algorithmic Impact Assessment: NHS Transparency Standard for Machine Learning Tools (UK Department of Health and Social Care 2023).
[147] World Health Organization, Ethics and Governance of Artificial Intelligence for Health (2021).
[148] FDA, Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Action Plan (2021) https://www.fda.gov/media/145022/download accessed 18 April 2025.
[149] Ziad Obermeyer et al, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' (2019) 366 Science 447.
[150] European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) COM (2021) 206 final.
[151] Patient Protection and Affordable Care Act 2010, 42 USC Sect. 18116.
[152] Alexander v Sandoval 532 US 275 (2001); see also Griggs v Duke Power Co 401 US 424 (1971).
[153] HHS OCR, Bulletin on Civil Rights and Artificial Intelligence (2023) https://www.hhs.gov/ocr accessed 18 April 2025.
[154] Matthew McCoy et al, 'Clinical Decision Support and Liability for Medical Injuries in the Age of AI' (2021) 49 Journal of Law, Medicine & Ethics 44.
[155] Helling v Carey 519 P 2d 981 (Wash 1974).
[156] Canterbury v Spence 464 F 2d 772 (DC Cir 1972).
[157] Regulation (EU) 2016/679 (General Data Protection Regulation), art 22.
[158] Brent Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (2019) 1 Nature Machine Intelligence 501.
[159] Joanna C. Schwartz, 'A Dose of Reality for Medical Malpractice Reform' (2014) 88 New York University Law Review 1224.
[160] 21USC Sec. 360c(a)(1); FDA, Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning-Based Software as a Medical Device (2019).
[161] Irene Y Chen et al, 'Ethical Machine Learning in Health Care' (2020) 372 New England Journal of Medicine 401.
[162] Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code (Polity Press 2019); Matthew DeCamp and Kayte Spector-Bagdady, 'Fairness and Justice in Medical Algorithms' (2021) 385 The Lancet 192.
[163] FDA, Digital Health Advisory Committee Charter (2023) https://www.fda.gov/media/171620/download accessed April 2025.
[164] European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) COM (2021) 206 final.
[165] 42 USC Sect. 1395 et seq.
[166] NYC Local Law 144 (2021); New York City Department of Consumer and Worker Protection, Rules on Automated

Employment Decision Tools (2023).
[167] Treasury Board of Canada Secretariat, Directive on Automated Decision-Making (2020).
[168] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap' (2018) 51 UC Davis Law Review 399, 444.
[169] Elizabeth Adams and Eleanor Dashiell, 'Algorithmic Immunity? AI and Clinical Standard of Care' (2023) 28 Harvard Journal of Law & Technology 182.
[170] 42 USC Sec. 300aa-10 to 300aa-34; see also Dov Greenbaum and Mark Gerstein, 'Protecting Patients from AI-Related Harm in Healthcare: Liability, Regulation, and Redress' (2022) 29 Journal of Law & the Biosciences 1.
[171] European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) COM (2021) 206 final.
[172] Andrea Renda, 'The EU Artificial Intelligence Act: Between Opportunity and Responsibility' (2022) 13 European Journal of Risk Regulation 1.
[173] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 76.
[174] UK National Institute for Health and Care Excellence (NICE), Evidence Standards Framework for Digital Health Technologies (2022); PBAC, Guidelines for Preparing Submissions (Australian Government, 2022).
[175] Shirin Heidari et al, 'Artificial Intelligence in Global Health: Defining a Collective Path Forward' (2022) 7 Nature Digital Medicine 75.
[176] World Health Organization, Ethics and Governance of Artificial Intelligence for Health (WHO 2021).
[177] Stephanie Russo Carroll et al, 'Operationalizing the CARE and FAIR Principles for Indigenous Data Futures' (2021) 5 Scientific Data 1.
[178] OECD, Recommendation of the Council on Artificial Intelligence (2019); UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021); Council of Europe, Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (Draft, 2023).
[179] Margaret Mitchell et al, 'Model Cards for Model Reporting' (2019) Proceedings of the Conference on Fairness, Accountability, and Transparency 220; IBM Research, 'FactSheets: Increasing Trust in AI Services' (2019).
Jonathan Patz JL & Tech 353.
[181] Shapiro J and Donelson A, 'Insuring Against Algorithmic Discrimination' (2023) Yale J Health Pol'y L & Ethics (forthcoming).
[182] W. Nicholson Price II, 'Black-Box Medicine' (2015) 28 Harv JL & Tech 419; see also Riegel v Medtronic Inc 552 US 312 (2008).
[183] Veena Dubal, 'Algorithmic Wage Discrimination: Data, Inequality, and the Law' (2021) 121 Columbia Law Review 47.
[184] Ontario Federation of Labour v Uber Canada [2023] ONSC 1892.
[185] Irene Y. Chen et al, 'Can AI Help Reduce Disparities in General Medical and Mental Health Care?' (2020) 2 AMA Journal of Ethics E778.
[186] Jonathan Patz et al, 'Climate Change and Global Health: Quantifying a Growing Ethical Crisis' (2020) 12 Lancet Planetary Health e140.
[187] White House Environmental Justice Advisory Council, 'Justice40 Initiative Interim Implementation Guidance' (2021); WHO and WMO, Climate Services for Health: Improving Public Health Decision-Making in a Changing Climate (2023).
[188] Sonia K Katyal, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66 UCLA L Rev 54; Frank Pasquale, New Laws of Robotics: Defending Human Expertise in the Age of AI (Harvard UP 2020).
[189] US Food and Drug Administration, 'Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices' (2024) https://www.fda.gov/medical-devices accessed 22 April 2025.
[190] Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor (St Martin's Press 2018).
[191] Proposed in this article, see Section 6.3.3; cf. National Childhood Vaccine Injury Act of 1986, 42 USC § 300aa–10.
[192] European Commission, Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) COM(2021) 206 final.
[193] OECD, 'OECD Principles on Artificial Intelligence' (2019); WHO, Ethics and Governance of Artificial Intelligence for Health (2021); UNESCO, Recommendation on the Ethics of Artificial Intelligence (2022).
[194] Irene Y Chen et al, 'This Looks Like That: Deep Learning for Interpretable Image Recognition' (2020) 33 NeurIPS 8930; Sandra Wachter, Brent Mittelstadt, and Chris Russell,
'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI' (2021) 43 Human Rights Law Review 65.
[195] Veena Dubal, 'The Afterlife of Algorithmic Management: Labour, Law, and Resistance in the Platform Economy' (2023) 28 Columbia Journal of Race and Law 17; Jonathan Patz et al, 'Climate Change and Health: A Call for Evidence-Based, Equitable Policies' (2022) 398 The Lancet 626.