2024

Decoding U.S. Tort Liability in Healthcare's Black-Box AI Era: Lessons from the European Union

Mindy Duffourc

Sara Gerke

Decoding U.S. Tort Liability in Healthcare's Black-Box AI Era: Lessons from the European Union

Mindy Duffourc[*] & Sara Gerke[**]

27 STAN. TECH. L. REV. 1 (2024)

ABSTRACT

The rapid development of sophisticated artificial intelligence ("AI") tools in healthcare presents new possibilities for improving medical treatment and general health. Currently, such AI tools can perform a wide range of health-related tasks, from specialized autonomous systems that diagnose diabetic retinopathy to general-use generative models like ChatGPT that answer users' health-related questions. On the other hand, significant liability concerns arise

---

[*] Assistant Professor of Law, (Maastricht) Law and Tech Lab, Maastricht European Private Law Institute, Maastricht University, Maastricht, Netherlands; Research Team Member, Validating AI in Classifying Cancer in Real-Time Surgery (CLASSICA), European Union (Grant Agreement no. 101057321), Penn State Dickinson Law, Carlisle, PA, USA.

[**] Assistant Professor of Law, Penn State Dickinson Law, Carlisle, PA, USA; Co-Principal Investigator, WP8 (Legal, Ethical & Liability), Validating AI in Classifying Cancer in Real-Time Surgery (CLASSICA), European Union (Grant Agreement no. 101057321); Co-Principal Investigator, WP4 (Addressing Ethical/Legal Concerns), Optimizing Colorectal Cancer Prevention through Personalized Treatment with Artificial Intelligence (OperA), European Union (Grant Agreement no. 101057099); Multiple Principal Investigator, Bioethical, Legal, and Anthropological Study of Technologies (BLAST), National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the National Institutes of Health Office of the Director (NIH OD) (Grant Agreement no. 1R21EB035474-01); Co-Investigator (Supplemental Project), PREMIERE: A PREdictive Model Index and Exchange Repository, NIBIB and NIH OD (Grant Agreement no. 3R01EB027650-03S1); Co-Investigator, Penn State TCORS: Tobacco Product Composition Effects on Toxicity and Addiction, National Institute on Drug Abuse (NIDA)/National Institutes of Health (NIH) (Grant Agreement no. 1U54DA058271-01).

This Article was selected for the AALS Section on Law, Medicine and Health Care's "New Voices in Law, Medicine and Health Care" Program at the 2024 AALS Annual Meeting. This Article greatly benefited from feedback from participants at the Texas A&M University School of Law's Annual Health Law Conference, the Regulation and Innovation in the Biosciences (RIBS) Workshop, and the AALS Annual Meeting. For helpful comments and conversations, we would like to thank Valarie Blake, Anjali Deshmukh, Wendy Netter Epstein, George Horvath, Nicole Huberfeld, Ryan Knox, Craig Konnoth, Matthew Lawrence, Myrisha S. Lewis, Brendan Maher, Elizabeth McCuskey, Govind Persad, Jessica Roberts, Christopher Robertson, David A. Simon, Michael Sinha, Charlotte Tschider, Allison Whelan, and Carleen Zubrzycki. We thank Kaci McNeave and Robin Platte for excellent research assistance. All errors are our own.

as medical professionals and consumers increasingly turn to AI for health information. This is particularly true for black-box AI because while potentially enhancing the AI's capability and accuracy, these systems also operate without transparency, making it difficult or even impossible to understand how they arrive at a particular result.

The current liability framework is not fully equipped to address the unique challenges posed by black-box AI's lack of transparency, leaving patients, consumers, healthcare providers, AI manufacturers, and policymakers unsure about who will be responsible for AI-caused medical injuries. Of course, the United States is not the only jurisdiction faced with a liability framework that is out of tune with the current realities of black-box AI technology in the health domain. The European Union has also been grappling with the challenges that black-box AI poses to traditional liability frameworks and recently proposed new liability Directives to overcome some of these challenges.

As the first to analyze and compare the liability frameworks governing medical injuries caused by black-box AI in the United States and European Union, this Article demystifies the structure and relevance of foreign law in this area to provide practical guidance to courts, litigators, and other stakeholders seeking to understand the application and limitations of current and newly proposed liability law in this domain. We reveal that remarkably similar principles will operate to govern liability for medical injuries caused by black-box AI and that, as a result, both jurisdictions face similar liability challenges. These similarities offer an opportunity for the United States to learn from the European Union's newly developed approach to governing liability for AI-caused injuries. In particular, we identify four valuable lessons from the European Union's approach. First, a broad approach to AI liability fails to provide solutions to some challenges posed by black-box AI in healthcare. Second, traditional concepts of human fault pose significant challenges in cases involving black-box AI. Third, product liability frameworks must consider the unique features of black-box AI. Fourth, evidentiary rules should address the difficulties that claimants will face in cases involving medical injuries caused by black-box AI.

TABLE OF CONTENTS

INTRODUCTION

> "*I do not intend to deceive or mislead anyone, but I sometimes make mistakes or assumptions based on incomplete or inaccurate data. I also do not have the clinical judgment or the ethical responsibility of a human doctor or nurse.*" – ChatGPT.[1]

In April 2023, OpenAI's ChatGPT-4 made headlines when it "passed the [United States] medical licensing exam with flying colors."[2] It also "impressed and horrified" a physician when it matched his medical expertise by correctly diagnosing his patient with a rare medical condition using information from the patient's medical record.[3] With more than 100 million users, ChatGPT's natural language responses to a wide range of inquiries and prompts have the potential to revolutionize the delivery of health information.[4]

In June 2023, a 29-year-old woman was awoken from her nap by her Apple Watch when an alarm sounded to alert her that her heart rate had risen to 178 and remained high for over ten minutes.[5] At the hospital, she was diagnosed with a blood clot in her lungs and underwent life-saving surgery.[6] The rise of wearable technology using artificial intelligence ("AI") in combination with new health applications brings enormous potential for consumers to prevent and manage health conditions.[7]

Healthcare providers and consumers alike are turning more frequently to AI systems and models to prevent, discover, monitor, and manage health conditions. For example, physicians may be tempted to turn to AI to answer clinical questions, provide medical information, or generate treatment

---

[1] Hilary Brueck, *The Newest Version of ChatGPT Passed the US Medical Licensing Exam with Flying Colors—and Diagnosed a 1 in 100,000 Condition in Seconds*, BUS. INSIDER (Apr. 6, 2023), https://perma.cc/K2MN-SGAY.
[2] *Id.*
[3] *Id.*
[4] *See* David Shepardson, *US Senate Leader Schumer Calls for AI Rules as ChatGPT Surges in Popularity*, REUTERS (Apr. 13, 2023, 8:18 PM UTC), https://perma.cc/TC2W-9E3J (noting that ChatGPT "has become the fastest-growing consumer application in history with more than 100 million monthly active users").
[5] Shubham Singh, *Apple Watch Alerts High Heart Rate, Saves 29-Year-Old Woman's Life from Fatal Blood Clot*, BUS. TODAY (June 20, 2023, 6:20 PM IST), https://perma.cc/RD9W-6MAE.
[6] *Id.*
[7] *See* Madelyn Knowles et al., *Consumer Adoption of Digital Health in 2022: Moving at the Speed of Trust*, ROCK HEALTH (Feb. 21, 2023), https://perma.cc/GTY7-HS5B (noting that wearable technology is approaching majority acceptance).

documentation.[8] Consumers, too, may try to save time and money by seeking quick answers to their medical questions and concerns from AI. In fact, a recent study found that when compared with human physicians, ChatGPT provided better quality and more empathetic responses to health questions posted in online fora.[9] In addition to generative AI like OpenAI's ChatGPT and Google's Bard, which are developed to answer questions on a variety of general topics (called here "general-use AI"), AI can also be designed specifically for healthcare ("health-specific AI"). For example, Microsoft Cloud for Healthcare, a suite of software designed to "[a]ccelerate innovation and improve healthcare experiences with AI-powered solutions," uses AI to help healthcare providers "make data-driven decisions."[10] AI can also be trained to interpret medical images, diagnose diseases, and assist with minimally invasive procedures.[11]

AI creates new possibilities for healthcare, including improving both access to care and patient outcomes, but it can also produce inaccurate medical information and recommendations, endangering patient safety. For example, the accuracy of an AI's output is highly dependent on the quality of its training data, which can be limited or biased. Even with high-quality training data, AI may generate recommendations or decisions based solely on the patterns and correlations it detects in medical data while failing to consider other factors

---

[8] *See* Geoff Brumfiel, *Doctors Are Drowning in Paperwork. Some Companies Claim AI Can Help*, NPR (Apr. 5, 2023, 5:13 AM ET), https://perma.cc/3BFD-DAMZ (discussing the development of AI technologies that can assist doctors with medical diagnosis, treatment plans, and documentation); Michael DePeau-Wilson, *GPT-4 Is Here. How Can Doctors Use Generative AI Now?*, MEDPAGE TODAY (Mar. 20, 2023), https://perma.cc/N9GD-DG53 (discussing the potential for generative AI to assist doctors with medical documentation and research).

[9] John W. Ayers et al., *Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum*, 183 JAMA INTERNAL MED. 589, 589 (2023). *But see* Douglas Johnson et al., *Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model*, NATURE PORTFOLIO (forthcoming) (manuscript at 6), https://perma.cc/ZD9H-JXYM (Feb. 28, 2023) ("While the AI-generated answers displayed high accuracy and completeness scores across various specialties, question types, and difficulty levels, further development is needed to improve the reliability and robustness of these tools before clinical integration.").

[10] *Microsoft Cloud for Healthcare*, MICROSOFT, https://perma.cc/2Z2G-6566; Tom McGuinness, *Microsoft Cloud for Healthcare: Empowering Healthcare to Deliver Meaningful Outcomes*, MICROSOFT (Apr. 12, 2023), https://perma.cc/NM4H-F2NZ.

[11] *See Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices*, U.S. FOOD & DRUG ADMIN., https://perma.cc/6DW2-F96H (Dec. 6, 2023); Daichi Kitaguchi et al., *Artificial Intelligence-Based Computer Vision in Surgery: Recent Advances and Future Perspectives*, 6 ANNALS GASTROENTEROLOGICAL SURGERY 29 (2022).

relevant to the patient's health.[12] An alarming issue with large language models ("LLMs"), like ChatGPT, is their tendency to "hallucinate" fabricated information and present it to the user with a high level of confidence, which can be misleading and dangerous.[13] The AI's ability to produce inaccurate outputs is particularly worrisome when it comes to "black-box" AI, which operates without transparency or the ability to explain the basis for its recommendations or decisions. This is because when users cannot understand how or why the AI arrived at a particular recommendation or decision, it becomes difficult, or even impossible, for them to recognize and correct potential errors in the AI's output.

As black-box AI becomes increasingly integrated into healthcare, U.S. courts will inevitably confront the task of examining complex liability issues that arise from medical injuries caused by such systems. The present tort framework governing liability for manufacturers and healthcare providers is often insufficient to address the unique challenges posed by black-box AI used in healthcare because it is not always possible to sufficiently link a black-box AI's injury-causing output to the actions of a legally responsible party. This is in contrast to transparent and interpretable "white-box" AI, which allows developers and users to understand the AI's reasoning process and evaluate its output, making it possible to trace injuries to the actions of manufacturers and/or healthcare providers under the existing liability framework.

Across the pond, the European Union has been grappling with similar challenges that AI poses to the existing frameworks that govern liability of manufacturers and healthcare providers. In an effort to address these liability challenges, the European Commission ("EC") recently proposed two Directives: a *Directive on Liability for Defective Products* ("Proposed Product Liability

---

[12] *See* Rich Caruana, et al., *Intelligible Models for Healthcare: Predicting Pneumonia Risk and Hospital 30-Day Readmission*, KDD '15: PROC. 21ST ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING, Aug. 10, 2015, at 1721 (noting that an AI system trained with medical records to triage patients with pneumonia confused correlation with causation in risk evaluation and found that patients with asthma were low risk because it associated better outcomes in those patients with their asthma condition rather than with the fact that they received a higher level of care because of their asthma diagnosis).

[13] *See* Brueck, *supra* note 1 (describing how ChatGPT can make mistakes and hallucinate information); Brumfiel, *supra* note 8 (noting that ChatGPT "can also hallucinate findings and fabricate sources" when diagnosing medical conditions).

Directive," or "Proposed PLD")[14] and a *Directive on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence* ("Proposed AI Liability Directive," or "Proposed AILD").[15] This Article argues that, given the similarities in the basic liability frameworks in the United States and the European Union and the shared liability challenges that black-box AI poses in cases involving medical injuries, the European Union's approach to AI liability can provide valuable insights as courts, litigators, and other stakeholders, including policymakers, manufacturers, and healthcare providers, in the United States begin to confront the challenges raised by black-box AI in healthcare.

This Article is the first to conduct an in-depth comparative analysis of the issue of liability for medical injuries caused by black-box AI in both the United States and the European Union to provide a practical guide for U.S. courts, AI manufacturers, healthcare providers, patients, or anyone else seeking to understand the application and limitations of the liability law in this domain. It consists of five Parts.

Part I provides an introduction to black-box AI in healthcare. It sketches the current state of AI in the health domain and explains the technical characteristics of various types of AI in an easily understandable fashion. It also provides specific examples of different types of AI currently in use and forecasts AI's future potential health-related uses.

Part II identifies and examines the current tort liability framework applicable to cases involving medical injuries caused by black-box AI in the United States. It discusses the legal elements required to hold manufacturers and healthcare providers liable in such cases and reveals that the existing framework struggles to govern some medical injuries caused by black-box AI. In particular, it exposes three main weaknesses in the current tort liability framework relating to: (1) legal standards used to determine product defects, (2) identification of fault by a legally responsible party, and (3) burden of proof rules.

---

[14] *Commission Proposal for a Directive of the European Parliament and of the Council on Liability for Defective Products*, COM (2022) 495 final (Sept. 28, 2022) [hereinafter Proposed PLD].

[15] *Commission Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive)*, COM (2022) 496 final (Sept. 28, 2022) [hereinafter Proposed AILD].

Part III identifies and analyzes the current and proposed liability framework applicable to cases involving medical injuries caused by black-box AI in the European Union. It first outlines the general liability principles that would govern manufacturer and healthcare provider liability for such injuries in the European Union under the current framework. It then provides an overview of the Proposed PLD and Proposed AILD and their potential impact on liability for medical injuries caused by black-box AI. Part III reveals that the European Union's basic liability framework uses legal principles remarkably similar to those in the United States by generally anchoring manufacturer liability on product defects and healthcare provider liability on fault. As a result, the European Union's newly proposed Directives, which are aimed at addressing the liability challenges caused by black-box AI, can provide important lessons as various stakeholders confront similar challenges in the United States.

Part IV aims to discover valuable insights from the European Union's approach by analyzing liability in three hypothetical scenarios involving a medical injury caused by black-box AI under both U.S. and EU law. The comparative analysis of these three scenarios reveals that liability gaps can manifest in both jurisdictions when a medical injury results from a healthcare provider's use of an autonomous black-box AI ("Scenario 1"), a non-autonomous black-box AI ("Scenario 2"), and a consumer's use of a general black-box generative AI ("Scenario 3"). These gaps are likely to occur when the AI's noninterpretable reasoning process produces an injury-causing output, and the following are true: (1) the AI was functioning as designed by the manufacturer, (2) the manufacturer complied with safety regulations relating to the AI's development and marketing, and (3) individual and organizational healthcare providers were reasonable in their selection and implementation of the AI and could not have reasonably known that the AI's output was incorrect. In addition, black-box AI's complex architecture and operation present evidentiary challenges for injured claimants.

Part V identifies lessons learned from the European Union's approach to AI liability in the context of medical injuries caused by black-box AI. It first highlights that the European Union's approach is relevant because the liability frameworks in both jurisdictions operate under similar legal principles and face common legal challenges when confronted with black-box AI in the healthcare domain. It then identifies four lessons learned from the European Union's approach that may help address these common challenges: (1) a broad approach to AI liability fails to provide solutions to some challenges posed by black-box AI in healthcare; (2) traditional concepts of human fault pose

significant challenges in cases involving black-box AI; (3) product liability frameworks must consider the unique features of black-box AI; and (4) evidentiary rules should address the difficulties that claimants will face in cases involving medical injuries caused by black-box AI.

I. BLACK-BOX AI IN HEALTHCARE

AI refers broadly "to the ability of machines to perform tasks that would normally require human-level intelligence."[16] While AI design and architecture can be complex and wide-ranging, one must first have a basic understanding of AI and its applications to engage in the liability analysis. Thus, this Part first explains the key characteristics of AI in healthcare and then discusses healthcare provider and direct-to-consumer uses.

A. Characteristics of AI in Healthcare

AI can be described as *autonomous* to mean that a human does not review or oversee their output, or *non-autonomous* to indicate that a human exercised some review or oversight of the AI's output.[17] AI can also be designed with varying levels of complexity. *White-box AI* is designed to be transparent and interpretable such that the developer and user are able to access and understand the logic underlying the AI's output. The simplest white-box AI receives an input and follows an explicit set of pre-programmed "if-then" rules to produce an output.[18] More complex white-box AI can be designed using *machine learning*, which involves training an AI with large amounts of data so that it can learn and improve its performance based on patterns and relationships in the data and without explicit rules.[19]

---

[16] Prathamesh Thakare, *The Future Is Now: A Look at the Most Promising Emerging Technologies*, SCRRUM (Jan. 1, 2023), https://perma.cc/Q3ZJ-FQUX.

[17] *See* Michael D. Abràmoff et al., *Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process*, 214 AM. J. OPHTHALMOLOGY 134, 134 (2020) (describing autonomous AI systems in healthcare as "AI systems that make clinical decisions without human oversight"). *But see* Mindy Nunez Duffourc, *Malpractice by the Autonomous AI Physician*, 2023 ILL. J.L. TECH. & POL'Y 1, 10–11 (2023) (noting a more dynamic approach to levels of automation and autonomy).

[18] KARTIK HOSANAGAR, A HUMAN'S GUIDE TO MACHINE INTELLIGENCE: HOW ALGORITHMS ARE SHAPING OUR LIVES AND HOW WE CAN STAY IN CONTROL 106–07 (2020).

[19] Oriol Pujol, *The Concept of "AI" Opacity and Societal Impact*, *in* ARTIFICIAL INTELLIGENCE AND THE LAW 28 (Pablo García Mexía & Francisco Pérez Bes eds., 2021).

Even more complex is AI that uses a subset of machine learning called *deep learning*. Deep learning enables AI to make decisions by training artificial neural networks with large datasets. Deep-learning AI is usually referred to as *black-box AI* because it uses complex algorithmic reasoning that makes it extremely difficult or impossible for humans to understand how the AI arrived at its output, making it noninterpretable.[20] Although computer scientists can explain the complex technical components used to design deep-learning AI, even they cannot truly understand exactly why a black-box AI produces a particular output.[21] Noninterpretability is especially concerning for "adaptive" black-box AI, which, instead of being "locked" when placed on the market to ensure that it produces the same output in response to the same input, can continue to learn as it collects new data while in use.[22] Generative AI is a type of black-box AI that uses deep-learning algorithms to recognize characteristics and patterns in vast amounts of data to create new content, such as text and images.[23]

Although the complexity of black-box AI makes it noninterpretable and often more unpredictable, this complexity also enables the AI to be more accurate and solve more complex problems.[24] Still, the combination of black-box AI's capacity to produce erroneous information and humans' inability to understand how it generated such outputs is a significant source of concern.

---

[20] Sara Gerke, *Health AI for Good Rather than Evil? The Need for a New Regulatory Framework for AI-Based Medical Devices*, 20 YALE J. HEALTH POL'Y L. & ETHICS 433, 441–42 (2021); *see also* Kun-Hsing Yu et al., *Artificial Intelligence in Healthcare*, 2 NATURE BIOMEDICINE ENG'G 719, 720 (2018).

[21] *See* Jason Yosinski et al., *Understanding Neural Networks Through Deep Visualization*, ARXIV (June 22, 2015, 12:57:15 UTC), https://perma.cc/8XAL-A2GU ("Neural networks have long been known as 'black boxes' because it is difficult to understand exactly how any particular, trained neural network functions due to the large number of interacting, non-linear parts. Large modern neural networks are even harder to study because of their size . . . .").

[22] U.S. FOOD & DRUG ADMIN., PROPOSED REGULATORY FRAMEWORK FOR MODIFICATIONS TO ARTIFICIAL INTELLIGENCE/MACHINE LEARNING (AI/ML)-BASED SOFTWARE AS A MEDICAL DEVICE (SAMD): DISCUSSION PAPER AND REQUEST FOR FEEDBACK 3 (2019), https://perma.cc/C3TJ-F56K [hereinafter FDA Discussion Paper].

[23] *Generative Models*, OPENAI (June 16, 2016), https://perma.cc/H29D-FUXW; *see also generally* Yosinski et al., *supra* note 21 ("Neural networks have long been known as 'black boxes' because it is difficult to understand exactly how any particular, trained neural network functions due to the large number of interacting, non-linear parts."); Joseph Rocca, *Understanding Variational Autoencoders (VAEs)*, TOWARDS DATA SCI. (Sept. 24, 2019), https://perma.cc/F5VS-3Q8B (describing the main types of deep-learning generative models).

[24] *See* Charlotte A. Tschider, *Medical Device Artificial Intelligence: The New Tort Frontier*, 46 BYU L. REV. 1551, 1561–62 (2021).

As a result, there are efforts to make black-box AI "explainable."[25] Explainable AI involves using a second explanatory white-box algorithm that can approximate the output of the black box and provide a *post hoc* explanation of the AI's output.[26] However, this explanation may not always be the true reason for the AI's output, which is often concealed from even the AI's creators.[27]

B. Healthcare Provider and Direct-to-Consumer Uses

Black-box AI is increasingly developed to be used by healthcare providers and consumers. This Section first discusses black-box AIs that are used by healthcare providers and then those that are directly addressed to consumers.

1. Healthcare Provider Use

Black-box AI can perform a wide range of healthcare tasks, from helping providers position patients for computed tomography ("CT") scans[28] to diagnosing medical conditions.[29] To date, the U.S. Food and Drug Administration ("FDA") has permitted the marketing of more than 690 AI-based medical devices, many of which are black-box AI that use deep-learning algorithms.[30] In addition to health-specific AI, healthcare providers may also turn to general-use AI, like ChatGPT, to obtain medical information.

---

[25] *See, e.g.*, Guang Yang et al., *Unbox the Black-Box for the Medical Explainable AI via Multi-Modal and Multi-Centre Data Fusion: A Mini-Review, Two Showcases and Beyond*, 77 INFO. FUSION 29 (2022).

[26] Boris Babic et al., *Beware Explanations from AI in Health Care*, 373 SCIENCE 284, 284 (2021); Sara Gerke, *"Nutrition Facts Labels" for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards*, 91 GEO. WASH. L. REV. 79, 90 (2023).

[27] *See* Babic et al., *supra* note 26, at 285 (warning that explainable AI may provide incorrect rationales for an AI decision); Gerke, *supra* note 26, at 112 (same).

[28] Yadong Gang, et al., *A Comparison Between Manual and Artificial Intelligence-Based Automatic Positioning in CT Imaging for COVID-19 Patients*, 31 EUR. RADIOLOGY 6049, 6057 (2021).

[29] *See* Huiying Liang, et al. *Evaluation and Accurate Diagnoses of Pediatric Diseases Using Artificial Intelligence*, 25 NATURE MED. 433, 433 (2019) (finding that black-box AI model diagnosed pediatric diseases with high accuracy).

[30] *See* U.S. FOOD & DRUG ADMIN., *supra* note 11 (listing devices that use deep learning AI models); Gerke, *supra* note 26, at 90 (noting that "[m]any AI/ML-based medical devices that are available on the U.S. market or in development, however, use deep learning").

a. Health-Specific AI

Currently, most legally-marketed black-box AIs are non-autonomous and intended to assist, but not replace, clinical decision making. Such devices can help radiologists interpret mammograms,[31] cardiologists evaluate coronary artery disease,[32] surgeons detect colorectal polyps,[33] and physicians assess musculoskeletal disease.[34]

Autonomous medical AI makes medical decisions without the involvement of a human medical expert.[35] In some cases, a black-box AI might have the technical ability to be autonomous but is classified as non-autonomous because the manufacturer's instructions require human expert involvement in the AI's use. For example, Viz ICH is a legally-marketed black-box AI that detects intracranial hemorrhages in CT scans and quickly notifies specialists of positive findings.[36] Viz ICH's analysis of the CT scans occurs in conjunction with the standard workflow involving human analysis of the same scans.[37] While using Viz ICH allows for quicker detection of intracranial hemorrhages and referral to a specialist, it does not replace the human radiologist's first-line analysis.[38] However, this same AI would be considered autonomous if it provided the sole first-line analysis of the CT scans. For example, in Europe, ChestLink was the

---

[31] *See* Letter from Thalia T. Mills, Dir. Div. of Radiological Health, U.S. Food & Drug Admin., to Juliane Dinter, Visage Imaging GmbH (Jan. 29, 2021), https://perma.cc/XQQ9-VV2Q.

[32] *See* Letter from Stephen C. Browning, Assistant Dir. Div. of Cardiac Electrophysiology, Diagnostics & Monitoring Devices, U.S. Food & Drug Admin., to Kelliann Payne, KeyaMed NA Inc. (Apr. 1, 2022), https://perma.cc/Q7HV-Y2HQ.

[33] *See* Letter from Shanil P. Haugen, Assistant Dir. Div. of Renal, Gastrointestinal, Obesity and Transplant Devices, U.S. Food & Drug Admin., to John Smith, Chengdu Wision Medical Device Co., Ltd. (Nov. 19, 2021), https://perma.cc/5BWZ-W93L.

[34] *See* Letter from Thalia T. Mills, Dir. Div. of Radiological Health, U.S. Food & Drug Admin., to M. Alaine Medio, Siemens Medical Solutions USA, Inc. (Mar. 16, 2020), https://perma.cc/DD7J-5C49.

[35] *See, e.g.*, Michael D. Abràmoff et al., *Pivotal Trial of an Autonomous AI-Based Diagnostic System for Detection of Diabetic Retinopathy in Primary Care Offices*, NPJ DIGIT. MED., Aug. 28, 2018, at 1–2 (describing an autonomous AI that interprets retinal images to detect diabetic retinopathy in patients without a human expert).

[36] Letter from Thalia T. Mills, Dir. Div. of Radiological Health, U.S. Food & Drug Admin., to Gregory Ramina, Viz.ai, Inc. (Mar. 18, 2020), https://perma.cc/X456-SY5U; *see also* Stavros Mtsoukas et al., *Pilot Deployment of Viz-Intracranial Hemorrhage for Intracranial Hemorrhage Detection: Real-World Performance in a Stroke Code Cohort*, 53 STROKE 418, 418–19 (2022), https://perma.cc/K6FM-R9QD.

[37] *See* Letter from Thalia T. Mills, *supra* note 36.

[38] *See id.* (noting that the algorithm analyzes images "in parallel to standard of care image interpretation").

first autonomous medical AI to receive a CE mark indicating compliance with the applicable EU requirements.[39] Unlike Viz ICH's analysis of CT scans, which are subsequently reviewed by human radiologists who are responsible for issuing the radiology reports, ChestLink uses deep learning to analyze chest X-rays and produce radiology reports without the involvement of a human radiologist in the first-line analysis.[40] These examples illustrate that AI is already capable of analyzing radiology films without human expert involvement.

In the United States, IDx-DR (rebranded as LumineticsCore) is a black-box autonomous medical AI that has been reviewed by the FDA and authorized for use in clinical practice.[41] IDx-DR uses deep learning to diagnose diabetic retinopathy, a sight-threatening eye condition, without input or oversight by a human ophthalmologist.[42] This allows healthcare providers without ophthalmological expertise to screen patients for diabetic retinopathy using IDx-DR, enabling early disease detection and management.[43]

As AI technology continues to advance and the growth of big data accelerates, the development and use of autonomous black-box AI is expected to increase in the health sector.[44] Autonomous deep-learning AI that can screen for hard-to-catch heart conditions,[45] analyze CT scans to provide early diagnosis of pleural diseases,[46] and analyze images of skin lesions to classify skin cancer[47]

---

[39] *See Autonomous AI Medical Imaging: Understanding ChestLink*, OXIPIT (Apr. 5, 2022), https://perma.cc/NLZ9-E53A (describing ChestLink's CE mark and autonomy).

[40] *See id.* (describing ChestLink's ability to autonomously produce final radiology reports when it "is highly confident that the image features no abnormalities").

[41] *See De Novo Classification Request for IDx-DR*, U.S. FOOD & DRUG ADMIN., https://perma.cc/H4FS-D33U; Press Release, Digit. Diagnostics, Digital Diagnostics Sheds Light on AI Tech with LumineticsCore™ Product Rebrand (Mar. 28, 2023), https://perma.cc/QZ4K-LUS2.

[42] *See* Press Release, U.S. Food & Drug Admin., FDA Permits Marketing of Artificial Intelligence-Based Device to Detect Certain Diabetes-Related Eye Problems (Apr. 11, 2018), https://perma.cc/6D6C-HP5F.

[43] *See id.*

[44] *See* Jordan Zheng Ting Sim et al., *Machine Learning in Medicine: What Clinicians Should Know*, 64 SING. MED. J. 91, 96 (2021).

[45] *See* Soonil Kwon et al., *Deep Learning Approaches to Detect Atrial Fibrillation Using Photoplethysmographic Signals: Algorithms Development Study*, 7 JMIR MHEALTH & UHEALTH, Dec. 6, 2019, at 2.

[46] *See* Neslihan Ozcelik et al., *Deep Learning for Diagnosis of Malign Pleural Effusion on Computed Tomography Images*, 78 CLINICS, Jan.–Dec. 2023, at 6.

[47] *See* Andre Esteva et al., *Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks*, 542 NATURE 115 (2017).

promises the potential to offer improved diagnostic timing and accuracy, less invasive medical procedures, increased access to treatment, and lower treatment costs.[48]

b. General-Use AI

Healthcare providers might also rely on general-use AI to help them with a variety of tasks, ranging from email drafting to diagnosis and treatment planning.[49] According to a physician and former CEO of Kaiser Permanente, "[n]o physician who practices high-quality medicine will do so without accessing ChatGPT or other forms of generative AI."[50] However, physicians' growing interest in using generative AI concerns bioethicists who "worry that doctors will turn to the bot for advice when they encounter a tough ethical decision."[51]

Nevertheless, there is evidence that doctors and medical students are already using ChatGPT to answer clinical questions, and when tested, ChatGPT was able to "produce a correct diagnosis accurately at close to the level of a third- or fourth-year medical student."[52] Of course, if ChatGPT provides accurate information, using this tool may indeed assist with a quick diagnosis. Indeed, ChatGPT has shown some promise in the area of medical diagnosis, but it also has important limitations, for example, omitting relevant differential diagnoses.[53] Finally, the confidence and sophistication with which ChatGPT

---

[48] *See, e.g.*, Sim et al., *supra* note 44, at 96 (describing the potential for AI to improve medical diagnosis and treatment); Walker Morrell et al., *The Oversight of Autonomous Artificial Intelligence: Lessons from Nurse Practitioners as Physician Extenders*, 9 J.L. & BIOSCIENCES 1, 5 (2022) (describing the potential for autonomous AI to increase access to healthcare and decrease the cost of care).

[49] *See* Brumfiel, *supra* note 8; Andrea Koncz, *6 Potential Medical Use Cases for ChatGPT*, MED. FUTURIST (Dec. 19, 2022), https://perma.cc/WK6H-PZRW (noting that ChatGPT may help doctors with drafting email, medical records, and research papers); *What Can ChatGPT Do for Healthcare Practices*, ENTREPRENEUR (May 2, 2023), https://perma.cc/M96M-ZC67 (noting that ChatGPT may help healthcare professionals with administrative tasks and patient communication).

[50] Khari Johnson, *ChatGPT Can Help Doctors—and Hurt Patients*, WIRED (Apr. 24, 2023, 7:00 AM), https://perma.cc/39AV-JSNM.

[51] *Id.*

[52] Brumfiel, *supra* note 8.

[53] *See* Stephen Hughes, *How Good Is ChatGPT at Diagnosing Disease? A Doctor Puts It Through Its Paces*, CONVERSATION (Apr. 27, 2023, 4:54 AM EDT), https://perma.cc/M267-X8VV.

responds to inquiries is troubling, given its ability to give fake information, often referred to as "hallucination."[54]

2. Direct-to-Consumer Use

Black-box AIs are also increasingly developed for and targeted to consumers (rather than healthcare providers). Direct-to-consumer ("DTC") AI tools can be health-specific, such as health apps or patient monitoring devices, but there are also general-use AIs like ChatGPT that consumers may consult for health information and advice.

a. Health-Specific AI

Consumers are more and more likely to turn to digital health resources like wearable technology and health apps for health-related information, management, and monitoring.[55] While the majority of health apps and wearables focus on general wellness and activity tracking, there is an increase in the development of DTC health technology aimed at managing health conditions, such as mental health disorders, diabetes, digestive diseases, and heart conditions.[56]

Many of these DTC digital health technologies are AI-driven.[57] Of these, some are black-box AI, including applications that can interpret at-home ECG's and provide users with "instant determinations of multiple cardiac

---

[54] *See* Jan Homolak, *Opportunities and Risks of ChatGPT in Medicine, Science, and Academic Publishing: A Modern Promethean Dilemma*, 64 CROAT. MED. J. 1, 2 (2023) ("The worst part is, the ability of ChatGPT to write a text of surprising quality might deceive reviewers and readers, with the final result being an accumulation of dangerous misinformation."); *see generally* Hussam Alkaissi & Samy I. McFarlane, *Artificial Hallucinations in ChatGPT: Implications in Scientific Writing*, CUREUS, Feb. 19, 2023 (describing how ChatGPT hallucinated medical sources).

[55] *See* Knowles, *supra* note 7 (noting the increased adoption of digital health technology); IQVIA, DIGITAL HEALTH TRENDS 2021, at 1 (2021), https://perma.cc/87DJ-RDD6 (noting more than 350,000 digital health apps were available to consumers as of 2021).

[56] IQVIA, *supra* note 55, at 8.

[57] Sara Gerke & Delaram Rezaeikhonakdar, *Privacy Aspects of Direct-to-Consumer Artificial Intelligence/Machine Learning Health Apps*, INTELLIGENCE-BASED MED., Apr. 7, 2022, at 1.

conditions"[58] and identify irregular heart rhythms.[59] The availability of DTC black-box AI for healthcare will only continue to grow as big data and machine learning technology advance. Google recently released an open-source software suite to facilitate the development of digital health apps.[60] In the future, consumers might use a health-specific LLM to answer medical questions[61] or deep-learning wearable technology to detect early stroke symptoms.[62]

b. General-Use AI

Consumers might also turn to general-use AI, like ChatGPT, for information and answers to questions that can have significant impacts on their health.[63] Consumers already rely on the internet to obtain health information and advice (the "Dr. Google" phenomenon).[64] However, unlike the simple retrieval of third-party information that consumers obtain using "Dr. Google," LLMs actively generate new content, which may be inaccurate or fabricated. Additionally, the level of sophistication and personalization of an LLM's responses might lull a consumer into thinking that the medical information it generates is more reliable than it actually is because "[p]eople tend to more naturally trust something that mimics human behaviors and responses, such as the responses generated by ChatGPT."[65] This false sense of reliability is further aggravated by the fact that it is difficult for a consumer, who is generally not a medical expert,

---

[58] *FDA Clears First of Its Kind Algorithm Suite for Personal ECG*, ALIVECOR (Nov. 23, 2020), https://perma.cc/TXF9-YGBL; *see also* IQVIA, *supra* note 55, at 26.

[59] *See* Letter from Jennifer Shih Kozen, Asst. Dir. Div. of Cardiac Electrophysiology, Diagnostics, & Monitoring Servs., U.S. Food & Drug Admin., to Luke Olson, Regul. Affs., Apple Inc. (June 3, 2022), https://perma.cc/9MYH-H2DL.

[60] Heather Landi, *Google Launches Open Health Stack for App Developers, Unveils New AI Partnerships*, FIERCE HEALTHCARE (Mar. 14, 2023, 1:00 PM), https://perma.cc/NZ9A-8W7G.

[61] *See id.* (noting that Google is investing in the development of medical LLMs).

[62] *See* Fei Jiang et al., *Artificial Intelligence in Healthcare: Past, Present, and Future*, 2 STROKE & VASCULAR NEUROLOGY 230, 240 (2017) (discussing the current uses and developments of complex AI models for stroke detection, diagnosis, and treatment).

[63] *See* Mindy Duffourc & Sara Gerke, *Generative AI in Health Care and Liability Risks for Physicians and Safety Concerns for Patients*, 330 JAMA 313, 313 (2023) (noting that patients might seek answers to their medical questions from ChatGPT).

[64] Joanna Burzyńska et al., *Dr. Google: Physicians—The Web—Patients Triangle: Digital Skills and Attitudes Towards e-Health Solutions Among Physicians in South Eastern Poland—A Cross-Sectional Study in a Pre-COVID-19 Era*, 20 INT'L J. ENV'T RSCH. & PUB. HEALTH 978, 978 (2023).

[65] Editorial, *Will ChatGPT Transform Healthcare?*, 29 NATURE MED. 505, 505 (2023).

to evaluate the accuracy and reliability of the medical information that ChatGPT provides.[66] One physician concluded that:

> I am most amazed by how clearly it communicates. But I am far less confident about how well it curates the information that it communicates. And that's the big problem that I see. Bill Gates noted that automation amplifies the efficiencies of current operations—but that it also amplifies the inefficiencies. Digital sources are already rife with medical misinformation, and my worry is that misinformation is more likely to be amplified by, rather than filtered out by, programs such as ChatGPT.[67]

There is little doubt that AI-driven health technology will increase in its development, adoption, and sophistication and bring with it many health-related benefits. These benefits are not without risks, however, and AI-driven healthcare presents liability risks when they cause medical injuries.

II. TORT LIABILITY FRAMEWORK FOR BLACK-BOX AI IN HEALTHCARE IN THE UNITED STATES

Currently, tort liability for medical injuries caused by black-box AI could be borne by manufacturers under product liability law and/or by individual and organizational healthcare providers under medical negligence law.[68] This Part first maps the current framework and exposes the challenges posed by black-box AI. Second, it considers the rules governing the burden of proof.

A. Manufacturer Liability

Manufacturers of black-box AI can face claims of strict or negligent product liability, though it is unclear whether AI would be considered a "product"

---

[66] *See* Johnson, *supra* note 50 ("Users also need to beware how ChatGPT-style bots can present fabricated, or 'hallucinated,' information in a superficially fluent way, potentially leading to serious errors if a person doesn't fact-check an algorithm's responses.").
[67] David A. Asch, *An Interview with ChatGPT About Health Care*, NEW ENG. J. MED. CATALYST (Apr. 4, 2023), https://perma.cc/XHZ3-E8T5.
[68] Contract liability based on breach of warranty is outside of the scope of this Article.

subject to product liability law.[69] There is some support for finding that software is a product because it is considered a "good" under the Uniform Commercial Code and taxed as a good.[70] Additionally, damage caused by information errors have been subject to product liability.[71] However, courts do not agree on whether data constitutes "tangible personal property," a defining characteristic of a "product" under the *Restatement (Third) of Torts: Products Liability*.[72]

Assuming that AI is considered a product, strict product liability would focus on the black-box AI's condition, whereas fault-based product liability focuses on the manufacturer's conduct. However, the legal analysis of strict and negligent product liability claims often overlaps because determining whether a product is defective requires some evaluation of the manufacturer's behavior relating to its design, manufacture, and marketing.[73] As a result, the general factors governing both strict and negligent product liability claims are similar.[74]

A black-box AI manufacturer might be strictly liable if:

(1) the AI was unreasonably dangerous (defective),

(2) when it left the manufacturer's control, and

(3) caused injury.[75]

---

[69] *See* Michael D. Scott, *Tort Liability for Vendors of Insecure Software: Has the Time Finally Come?*, 67 MD. L. REV. 425, 461–67 (2008) (noting that software might not be subject to product liability law).

[70] *Id.* at 462.

[71] *Id.* at 465–66.

[72] *Id.* at 463.

[73] *See* Denny v. Ford Motor Co., 662 N.E.2d 730, 735 (N.Y. 1995) (noting that the "risk/utility" test for strict product liability judges "the reasonableness of an actor's conduct . . . in light of a number of situational and policy-driven factors").

[74] *See* Kristine Cordier Karnezis, Annotation, *Products Liability: Modern Cases Determining Whether Product Is Defectively Designed,* 96 A.L.R.3d 22 § 2[a] (1979) (noting that negligent and strict liability for design defects requires a similar analysis). In fact, several states have codified product liability law to merge strict and fault-based liability principles. *See, e.g.*, LA. STAT. ANN. § 9:2800.52 (2023); MISS. CODE ANN. § 11-1-63 (West 2023); CONN. GEN. STAT. ANN. §§ 52-572m to -572n (West 2023); IND. CODE ANN. § 34-20-1-1 (West 2023); N.J. STAT. ANN. §§ 2A:58C-1 to -11 (West 2022); Kansas Products Liability Act, KAN. STAT. ANN. §§ 60-3301 to -3307 (West 2023); OHIO REV. CODE ANN. § 2307.71 (West 2023).

[75] *See* RESTATEMENT (SECOND) OF TORTS § 402A (AM. L. INST. 1965); Burlington N. & Santa Fe R.R. Co. v. ABC-NACO, 906 N.E.2d 83, 96 (Ill. App. Ct. 2009) (noting that a plaintiff must demonstrate "that the defective condition proximately caused the plaintiff's injuries" to prevail on a claim of strict product liability).

Negligent product liability requires the plaintiff to establish that "the defective condition of the product was the result of negligence."[76] A defective condition can result from at least one of three defects:

(1) design defect,

(2) manufacturing defect, and/or

(3) marketing defect related to the AI's instructions and warnings.[77]

The following analyzes these three defects and potential defenses available to black-box AI manufacturers.

1. Design Defects

Design defects stem from flaws in the design that make a product dangerous for its intended or reasonably foreseeable use.[78] A black-box AI's design may be defective if it presents an unreasonable danger to the consumer[79] and/or is not safe for its intended or reasonably foreseeable use.[80] Parties that partially contribute to the creation of a black-box AI, for example through writing portions of code or determining training data, might be considered component part manufacturers, who can be liable for injuries caused by defectively designed parts that were integrated into a final AI.[81]

---

[76] Niemela v. Imperial Mfg., Inc., 263 P.3d 1191, 1199 (Utah Ct. App. 2011) (emphasis omitted) (quoting Bishop v. GenTec Inc., 48 P.3d 218, 226 (Utah 2002)); *see also* Chestnut v. Ford Motor Co., 445 F.2d 967, 969 (4th Cir. 1971) (noting that a plaintiff must show that "the defect was the result of the defendant's failure to exercise due care" to prevail on a claim of negligent product liability).

[77] *See* Brown v. Super. Ct. of S.F., 751 P.2d 470, 475 (Cal. 1988) (noting three types of product defects: manufacturing, design, and instruction and warning). The Restatement (Third) of Torts also recognizes product liability for drugs and medical devices that have manufacturing or design defects or inadequate instructions for warnings. RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 6 (AM. L. INST. 1998).

[78] *See* RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 2 (AM. L. INST. 1998) ("A product . . . is defective in design when the foreseeable risk of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design by the seller . . . and the omission of the alternative design renders the product not reasonably safe . . . .").

[79] *See* Karnezis, *supra* note 74, § 3 (noting cases from Colorado, Florida, Indiana, Kentucky, Louisiana, Massachusetts, Montana, New Hampshire, New Jersey, New Mexico, Oklahoma, Rhode Island, South Carolina, Texas, Utah, Wisconsin).

[80] *See id.* § 4 (noting cases from Arizona, Arkansas, Florida, Georgia, Idaho, Illinois, Iowa, Kansas, Minnesota, Mississippi, Missouri, Nebraska, Nevada, New Jersey, New York, North Dakota, Ohio, Pennsylvania, South Dakota, Tennessee); *id.* § 5 (noting cases from California).

[81] *See generally* AM. L. PROD. LIAB. 3D § 8:14 (2023) (discussing potential liability for component part manufacturers).

However, component part manufacturers can escape such liability if they followed the design specifications of the end-product manufacturer and did not know or had no reason to know that the AI's design was defective.[82]

An unreasonably dangerous black-box AI will subject its manufacturer to liability for injuries it causes.[83] The *Restatement (Second) of Torts* describes an unreasonable danger as one that is "beyond that which would be contemplated by the ordinary consumer who purchases it, with the ordinary knowledge common to the community as to its characteristics."[84] However, as noted by the Connecticut Supreme Court, some products "involve[] complex product designs in which an ordinary consumer may not be able to form expectations of safety."[85]

The *Restatement (Third) of Torts: Products Liability* replaces the "consumer expectations" test with a "risk-utility" test, which weighs a product's risks and benefits to measure whether a product is unreasonably dangerous, including whether there was a reasonably alternative design that existed when the product was placed on the market.[86] The overwhelming majority of U.S. courts will likely incorporate some risk-utility considerations to determine whether a black-box AI is unreasonably dangerous in design defect claims.[87] Because the

---

[82] *See* 6 STUART M. SPEISER ET AL., AMERICAN LAW OF TORTS § 18:135 (2023) ("[T]he manufacturer of a component part of an unassembled product is not liable on the basis of design defect . . . where the component was manufactured in accordance with the design, plans, and specifications of the finished product manufacturer, so long as the plans and specifications do not reveal any inherent or obvious danger in either the component or in the assembled unit.").

[83] *See* Allan E. Korpela, Annotation, *Products Liability: Product as Unreasonably Dangerous or Unsafe Under Doctrine of Strict Liability in Tort*, 54 A.L.R.3d 352 § 2[a] (1973) (noting that in accordance with § 402A of the Second Restatement of Torts, "in most jurisdictions, for the doctrine to apply, the product must not only be in a defective condition, but such condition must be dangerous to a degree described as 'unreasonable'"). The Restatement (Third) of Torts employs a similar strict product liability for defective drugs and medical devices that are "not reasonably safe." RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 6 (AM. L. INST. 1998).

[84] RESTATEMENT (SECOND) OF TORTS § 402A cmt. a (AM. L. INST. 1965).

[85] Potter v. Chi. Pneumatic Tool Co., 694 A.2d 1319, 1333 (Conn. 1997). The court noted that in such cases, the consumer expectations test may be modified to include risk-utility factors. *Id.*

[86] RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 2 (AM. L. INST. 1998). *But see* Linda A. Sharp, Annotation, *Products Liability: Consumer Expectations Test*,73 A.L.R.5th 75 § 2[b] (1999) (noting that courts are reluctant to replace the Second Restatement's "consumer expectations" test with the Third Restatement's "risk-utility" test).

[87] *See* Branham v. Ford Motor Co., 701 S.E.2d 5, 14 n.11 (S.C. 2010) ("By our count 35 of the 46 states that recognize strict products liability utilize some form of risk-utility analysis in their approach to determine whether a product is defectively designed.").

complexity of black-box AI may exceed consumer's ability to understand and predict the safety risks of using these systems to obtain health-related information, the risk-utility test may be a better measure of product safety than the "consumer expectations" test.

The following factors can guide the risk-utility analysis: (1) the AI's usefulness; (2) the probability of danger; (3) the avoidability of danger; (4) the state of the art at the time the AI was manufactured; and (5) the effect of eliminating danger on the cost or utility of the AI.[88] Similarly, a negligent design claim would consider, "(1) the extent that the manufacturer could foresee that its actions would cause harm; (2) the likelihood of injury; (3) the magnitude of the burden of guarding against it; and (4) the consequences of placing the burden on the defendant."[89]

2. Manufacturing Defects

A black-box AI might suffer from a manufacturing defect if flaws in the manufacturing process depart from the intended design.[90] An AI with a manufacturing defect would be in a "substandard condition," making it more dangerous than anticipated by its design.[91] Generally, flaws in the manufacturing process can stem from poor construction, poor workmanship, or defective materials, which would lead to liability for the manufacturer.[92]

---

[88] *See In re* Cook Med., Inc., IVC Filters Mktg., Sales Pracs. & Prod. Liab. Litig., 431 F. Supp. 3d 1033, 1047 (S.D. Ind. 2020) (listing the factors considered in the risk-utility analysis).

[89] Slisze v. Stanley-Bostitch, 979 P.2d 317, 320 (Utah 1999).

[90] *See* Am. Tobacco Co. v. Grinnell, 951 S.W.2d 420, 434 (Tex. 1997) ("[A] plaintiff has a manufacturing defect claim when a finished product deviates, in terms of its construction or quality, from the specifications or planned output in a manner that renders it unreasonably dangerous."); RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 2(a) (AM. L. INST. 1998) ("A product . . . contains a manufacturing defect when the product departs from its intended design.").

[91] *See* Gall v. Smith & Nephew, Inc., 286 Cal. Rptr. 3d 108, 114 (Cal. Ct. App. 2021) (holding that there was no manufacturing defect because the defendant failed to show that the product left "the factory in substandard condition"); Longo v. E.I. Dupont De Nemours & Co., 632 So. 2d 1193, 1197 (La. Ct. App. 1994) (holding that there was no manufacturing defect because the product was "normally manufactured"); MacPherson v. Buick Motor Co., 111 N.E. 1050, 1053 (N.Y. 1916) (holding that the manufacturer was "responsible for the finished product").

[92] *See* Tears v. Bos. Sci. Corp., 344 F. Supp. 3d 500, 510 (S.D.N.Y. 2018) (describing how defects can be "a result of some mishap in the manufacturing process").

For black-box AI, this might relate to defects in the manufacturing of physical components of hardware used alongside the AI or it might include a less tangible part of the manufacturing process such as providing a training data set that departs from the training data set intended by its original design. Manufacturers of component parts of a black-box AI may be liable to the end user if (1) a component part suffered from a manufacturing defect and (2) the part did not undergo a "substantial change" when incorporated into the final AI.[93] The manifestation of a risk inherent from a black-box AI that was manufactured in accordance with its design specifications is not evidence of a manufacturing defect.[94]

3. Marketing Defects

A black-box AI manufacturer might be liable for injuries caused by inadequate instructions for the AI's use or inadequate warnings about potential risks of using the AI. Manufacturers usually have a duty to warn consumers of the inherent risks of which they knew or should have known,[95] including risks inherent in the intended use of the product as well as those inherent in reasonably foreseeable misuse.[96] Foreseeability concerns expectations that are objectively reasonable and not merely possible.[97]

Black-box AI manufacturers must also conduct post-market monitoring by reviewing "research, adverse reaction reports, scientific literature and other available" data about the AI to ensure that the existing warnings and instructions are sufficient.[98] Component part manufacturers have a duty to

---

[93] *See* SPEISER ET AL., *supra* note 82, § 18:135 (describing the conditions under which component parts manufacturers may be liable).

[94] *See Gall*, 286 Cal. Rptr. 3d at 114 (finding that the risk of a pseudotumor was "consistent with a perfect implant and is not probative of a defect").

[95] *See* Thacker v. Ethicon, Inc., 47 F.4th 451, 459 (6th Cir. 2022) (describing the elements of a failure to warn claim).

[96] *See* Mack v. Stryker Corp., 748 F.3d 845, 849 (8th Cir. 2014) (noting that a manufacturer has a duty to avoid unreasonable risk of harm for either an intended or unintended, yet foreseeable, use).

[97] *See, e.g.*, Winnett v. Winnett, 310 N.E.2d 1, 5 (Ill. 1974) (finding that it was not reasonably foreseeable that a child would approach and put their fingers into a piece of farm equipment while in operation).

[98] Rosen v. St. Jude Med., Inc., 41 F. Supp. 3d 170, 184 (N.D.N.Y. 2014); *see also* Baker v. St. Agnes Hosp., 421 N.Y.S.2d 81, 85 (N.Y. App. Div. 1979) (describing the "continuing obligation" of manufacturers); McLaughlin v. Bayer Corp., 172 F. Supp. 3d 804, 820 (E.D. Pa. 2016) (recognizing a claim for "negligent risk management").

warn the buyers that incorporate their component of inherent risks, but they generally do not have a duty to warn the ultimate consumers about the dangers of the final AI, unless they knew or should have known that the component would become unreasonably dangerous when integrated into the manufacturer's black-box AI design.[99]

4. Defenses

Black-box AI manufacturers can assert a variety of defenses in product liability claims. First, these claims may be preempted under section 521 of the Federal Food, Drug, and Cosmetic Act ("FDCA").[100] An AI with software functions that relate to "the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition" will likely be considered an AI-based medical device when the healthcare provider is *not* able to "independently review the basis" for the AI's output.[101] However, the medical device definition may not include some black-box AIs, such as those with functions intended for "maintaining or encouraging a healthy lifestyle."[102] Additionally, while manufacturers of AI-based medical devices that underwent the FDA's premarket approval ("PMA") will likely be entitled to preemption, preemption

---

[99] *See* RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 5 (AM. L. INST. 1998) (describing how component part manufacturers may be subject to liability for an injury caused by the final product); Levine v. Wyeth Inc., 684 F. Supp. 2d 1338, 1346 (M.D. Fla. 2010) (listing the elements for a products liability claim against a component part manufacturer); Longo v. E.I. Dupont De Nemours & Co., 632 So. 2d 1193, 1197 (La. Ct. App. 1994) (holding that component part manufacturer was not liable for the plaintiff's injury because the part "was not unreasonably dangerous"); Westphal v. E.I. du Pont de Nemours & Co., Inc., 531 N.W.2d 386, 393 (Wis. Ct. App. 1995) (holding that the component part manufacturer was not liable because "it would be unreasonable and impractical to require . . . [the manufacturer] to have intimate knowledge of all uses of all products made with [the component part]"); AM. L. PROD. LIAB. 3D, *supra* note 81, § 8:18.

[100] *See* Federal Food, Drug, and Cosmetic Act (FDCA) § 521, 21 U.S.C. § 360k (providing for preemption with regard to certain medical devices).

[101] FDCA § 520(o)(1)(B), 21 U.S.C. § 360j(o)(1)(B); *see* FDCA § 201(h)(1), 21 U.S.C. § 321(h)(1) (defining the term "device"); FDCA § 520(o)(1)(E), 21 U.S.C. § 360j(o)(1)(E) (excluding certain clinical decision support software functions from the definition of "device").

[102] *See* FDCA § 201(h)(1), 21 U.S.C. § 321(h)(1) (defining "device"); FDCA § 520(o)(1)(B), 21 U.S.C. § 360j(o)(1)(B) (excluding software functions intended "for maintaining or encouraging a healthy lifestyle" from the definition of "device"). For more information, see Gerke, *supra* note 26, at 96–121.

will likely be unavailable to those that underwent the less rigorous 510(k) process.[103]

Second, black-box AI manufacturers may also avoid strict product liability in design defect claims involving "unavoidably unsafe" AI.[104] An AI may be "unavoidably unsafe" if (1) at the time it was made, "it could not be made safe for its intended use even applying the best available testing and research; and (2) the benefits of the [product] justified its risk."[105] Drugs and medical devices have both been recognized as unavoidably unsafe products that justify an exemption because imposing liability "would be 'against the public interest' . . . because of 'the very serious tendency to stifle medical research and testing.'"[106] Black-box AI used to treat medical conditions or alleviate suffering may receive similar treatment as an "unavoidably unsafe" product, making it possible for manufacturers that do not undergo the PMA process to still avoid liability.[107]

Third, black-box AI manufacturers may also avoid product liability if the AI "conformed to the state of the art in existence at the time the product was . . . manufactured."[108] State of the art is usually determined by what is feasible to reflect what reasonably *can be done*, considering factors such as whether a design was practical, technologically and mechanically sound, and in

---

[103] *Compare* Riegel v. Medtronic, Inc., 552 U.S. 312 (2008) (preemption for devices that receive PMA), *with* Medtronic, Inc. v. Lohr, 518 U.S. 470 (1996) (no preemption for device that entered the market through 510(k) pathway). The vast majority of AI-based medical devices have so far gone through the 510(k) pathway. *See* U.S. FOOD & DRUG ADMIN, *supra* note 11; Gerke, *supra* note 26, at 131. For more information on FDA preemption, see W. Nicholson Price II et al., *Liability for Use of Artificial Intelligence in Medicine*, *in* RESEARCH HANDBOOK ON HEALTH, AI AND THE LAW (Barry Solaiman & I. Glenn Cohen eds., Edward Elgar Publ'g Ltd. forthcoming July 2024) (manuscript at 15–20), https://perma.cc/NQP4-JWCV.

[104] *See* RESTATEMENT (SECOND) OF TORTS § 402A cmt. k (AM. L. INST. 1965) (noting that manufacturers of products that are "incapable of being made safe for their intended and ordinary use" may avoid liability for "unfortunate consequences" of the product's use); Allen v. G.D. Searle & Co., 708 F. Supp. 1142, 1149 (D. Or. 1989) (noting that Oregon and "[m]ost other states have adopted section 402A and its comments").

[105] Burningham v. Wright Med. Tech., Inc., 448 P.3d 1283, 1291 (Utah 2019) (alteration in original).

[106] Brown v. Super. Ct. of S.F., 751 P.2d 470, 475 (Cal. 1988); *see also* Hufft v. Horowitz, 5 Cal. Rptr. 2d 377, 384 (Cal. Ct. App. 1992) (extending the *Brown* standard to medical devices).

[107] *See Burningham*, 448 P.3d at 1290–91 (recognizing affirmative defense based on comment k).

[108] Hughes v. Massey-Ferguson, Inc., 522 N.W.2d 294, 295 (Iowa 1994) (quoting IOWA CODE § 668.12 (1991)); *see also* Potter v. Chi. Pneumatic Tool Co., 694 A.2d 1319, 1346 (Conn. 1997) (summarizing the different approaches to state-of-the-art evidence).

compliance with government regulations.[109] While industry custom, which refers to what *is done* in the industry, can be used to establish the relevant state of the art in the industry, compliance with custom itself "does not establish conclusively the state of the art defense."[110]

Fourth, black-box AI manufacturers may also avoid liability under the learned intermediary doctrine ("LID"), which is recognized by many U.S. jurisdictions.[111] Under the LID, manufacturers that properly warn physicians of risks inherent in using the AI might not be liable for failing to warn patients who suffer injuries when such risks materialize.[112] This is because the LID essentially transfers the manufacturer's duty to warn the patient of the risks associated with the use of prescription drugs and medical devices to the physician.[113] The physician, as a "learned intermediary," must then use their independent medical judgment to assess and warn their patients of the dangers associated with the drugs and medical devices they prescribe.[114] However, the LID may not excuse black-box AI manufacturers from liability for failing to properly warn hospitals that purchase the AI systems even if they warn the physicians who use the AI because "the doctor is not a learned intermediary between the manufacturer and the hospital."[115]

---

[109] *See Hughes*, 522 N.W.2d at 295; Carter v. Massey-Ferguson, Inc., 716 F.2d 344, 347 (5th Cir. 1983); Montgomery Ward & Co. v. Gregg, 554 N.E.2d 1145, 1155–56 (Ind. Ct. App. 1990); Boatland of Hous., Inc. v. Bailey, 609 S.W.2d 743, 748 (Tex. 1980); Crittenden v. Fibreboard Corp., 794 P.2d 554, 555 n.2 (Wash. Ct. App. 1990); Frazier v. Kysor Indus. Corp., 607 P.2d 1296, 1301 (Colo. App. 1979), *rev'd on other grounds*, 642 P.2d 908 (Colo. 1982) (describing the state-of-the-art defense).

[110] *Hughes*, 522 N.W.2d at 296. The court reiterated that "[c]ustom refers to what was being done in the industry; state of the art refers to what feasibly could have been done." *Id.* at 295 (emphasis omitted) (quoting Chown v. USM Corp., 297 N.W.2d 218, 221 (Iowa 1980)).

[111] *See* RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 6 cmts. b, e (AM. L. INST. 1998) (discussing the LID); Dearinger v. Eli Lilly & Co., 510 P.3d 326, 329 (Wash. 2022) ("Every state in the country, along with the District of Columbia and Puerto Rico, has adopted the learned intermediary doctrine in some iteration.").

[112] RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 6 cmt. a (AM. L. INST. 1998).

[113] *See* Ellis v. C.R. Bard, Inc., 311 F.3d 1272, 1279–83 (11th Cir. 2002) (discussing the LID and its application for both prescriptions drugs and medical devices).

[114] *See id.*

[115] Taylor v. Intuitive Surgical, Inc., 389 P.3d 517, 525 (Wash. 2017). There, the court found that the manufacturer of the "da Vinci System," which is used to perform robotic surgery, could not rely on the LID to avoid liability for failing to warn the hospital even though it provided warnings to surgeons using the system. *Id.*

Finally, black-box AI manufacturers may also avoid liability when an injury resulted from a consumer's misuse, alteration, or modification of the AI, or from consumer negligence.[116] A consumer misuses a product when they "use[] the product in direct contravention of the product's warnings and instructions."[117] However, the manufacturer would still be liable if the consumer's misuse was reasonably foreseeable at the time the AI was sold.[118] Additionally, if the AI was altered or modified after it left the manufacturer's control, the manufacturer would generally not be liable for injuries caused by those alterations or modifications, unless they failed to warn of dangers stemming from foreseeable alterations or modifications.[119] Similarly, manufacturers may not be liable for injuries if the user was either contributorily negligent or knowingly assumed the risks that caused their injury.[120]

B. Healthcare Provider Liability

Individual and organizational healthcare providers that employ black-box AI can be directly liable for a patient's injury caused by their negligence.[121] They can also be held vicariously liable for injuries caused by the

---

[116] *See* Morgen v. Ford Motor Co., 797 N.E.2d 1146, 1149 (Ind. 2003); Campbell Hausfeld/Scott Fetzer Co. v. Johnson, 109 N.E.3d 953, 959 (Ind. 2018) (allowing a consumer's misuse to completely bar recovery); Hines v. Joy Mfg. Co., 850 F.2d 1146, 1146 (6th Cir.1988) (finding no liability for alterations and modifications). *But see* Mulherin v. Ingersoll-Rand Co., 628 P.2d 1301, 1304 (Utah 1981) (applying comparative fault principles to reduce recovery).

[117] *Campbell*, 109 N.E.3d at 959.

[118] *See Morgen*, 797 N.E.2d at 1150; *Campbell*, 109 N.E.3d at 959.

[119] *See* N.C. Gen. Stat. Ann. § 99B-3 (West); *Hines*, 850 F.2d at 1146 (no liability for alterations and modifications); Witthauer v. Burkhart Roentgen, Inc., 467 N.W.2d 439, 445 (N.D. 1991) (liability if failure to warn of foreseeable alternations and modifications).

[120] *See* Kupetz v. Deere & Co., 644 A.2d 1213, 1220 (Pa. Super. Ct. 1994) (noting that a consumer assumes a risk when they subjectively "know[] of the specific defect eventually causing [their] injury and voluntarily proceed[] to use the product with knowledge of the danger caused by the defect" (citing Berkebile v. Brantly Helicopter Corp., 337 A.2d 893, 901 (Pa. 1975))). *Compare* Smith v. Ingersoll-Rand Co., 14 P.3d 990, 996 (Alaska 2000) (noting that comparative fault is allowed in product liability actions), *with* Seay v. Chrysler Corp., 609 P.2d 1382, 1384 (Wash. 1980) (finding that comparative negligence is not applicable in strict product liability cases), *and* Mauch v. Mfrs. Sales & Serv., Inc., 345 N.W.2d 338, 347 (N.D. 1984) (finding that assumption of risk, but not comparative fault, can be a defense in a strict product liability claim).

[121] *See* Morrison v. MacNamara, 407 A.2d 555, 560 (D.C. 1979) ("The elements which govern ordinary negligence actions are also applicable in actions for professional negligence."); Elam v. Coll. Park Hosp., 183 Cal. Rptr. 156, 159 (Cal. Ct. App. 1982) (discussing the elements

negligence of a third party acting under their supervision and control.[122] However, a black-box AI itself cannot be negligent for its reasoning or output because it is not a legal person held to behavioral standards in tort.[123] To recover for an AI-caused injury under a negligence-based tort, a plaintiff must prove that the healthcare provider (1) owed them a duty of care and (2) breached that duty, which (3) caused (4) damages.[124] The following analyzes the first three elements of a patient's claim (i.e., duty, breach, and causation) and potential defenses of individual and organizational healthcare providers using black-box AI.

1. Duty

Individual healthcare providers that use black-box AI to treat patients have a duty to use the "degree of reasonable care and skill expected of members of the medical profession under the same or similar circumstances,"[125] taking into

---

of a medical malpractice claim in the United States); 3 SUMMARY PA. JURIS. 2D *Torts* § 37:45, Westlaw (database updated Oct. 2023) (discussing theories of direct hospital liability); *see also, e.g.*, W. Nicholson Price II, *Medical Malpractice and Black-Box Medicine*, *in* BIG DATA, HEALTH LAW, AND BIOETHICS 295, 299–304 (I. Glenn Cohen et al. eds., 2018) (discussing medical malpractice liability against providers and healthcare enterprises); A. Michael Froomkin et al., *When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning*, 61 ARIZ. L. REV. 33 (2019) (discussing medical liability rules in cases where machine learning systems become superior to human doctors); W. Nicholson Price II et al., *Potential Liability for Physicians Using Artificial Intelligence*, 322 JAMA 1765 (2019) (discussing generally when physicians might be held liable when using medical AI); Andrew D. Selbst, *Negligence and AI's Human Users*, 100 B.U. L. REV. 1315 (2020) (discussing how AI challenges negligence in general); W. Nicholson Price II et al., *How Much Can Potential Jurors Tell Us about Liability for Medical AI?,* 62 J. NUCLEAR MED. 15 (2021) (discussing liability in medical AI generally from the perspective of potential jurors).

[122] *See e.g.*, W.R. Habeeb, Annotation, *Liability of One Physician or Surgeon for Malpractice of Another*, 85 A.L.R.2d 889 § 4 (1962) (discussing vicarious liability for physicians); Cassandra P. Priestley, *Hospital Liability for the Negligence of Independent Contractors: A Summary of Trends*, 50 J. MO. BAR 263, 264 (1994) (discussing vicarious liability for hospitals); George Maliha et al., *Artificial Intelligence and Liability in Medicine: Balancing Safety and Innovation*, 99 MILBANK Q. 629, 631, 633 (2021) (discussing the AI ecosystem, including vicarious liability).

[123] Duffourc, *supra* note 17, at 18.

[124] *See Elam*, 183 Cal. Rptr. at 159 (listing elements of negligence in claim against healthcare organization); Jennings v. Badgett, 230 P.3d 861, 865 (Okla. 2010) (listing elements of negligence in claims against individual healthcare provider).

[125] *Morrison*, 407 A.2d at 561.

account the provider's training and specialization.[126] Organizational healthcare providers also owe a direct duty of care to patients to be reasonable in their use of black-box AI, including in the selection, implementation, and maintenance of the AI, training and oversight of employees that use AI, and the implementation of protocols and procedures associated with using AI.[127]

2. Breach

Individual healthcare providers are negligent when they breach the applicable standard of care by failing to "exercise the amount of care, skill, and diligence exercised generally in the community by doctors engaged in the same field."[128] The standard of care is flexible and can "depend[] on many factors, including a doctor's specialty, the resources available, and the advances of the medical profession at the time of the alleged negligent act."[129] Complicating the landscape for new medical technologies, including AI, is the fact that there is not always a single course of action required by the standard of care.[130] As a result, individual providers that use black-box AI to treat patients will generally not be negligent as long as their medical treatment is supported by "a considerable number of [their] professional brethren in good standing in [their] community."[131]

3. Causation

If a healthcare provider who uses black-box AI to treat a patient breaches the standard of care, this breach will sufficiently cause a patient injury when it

---

[126] *See* Betesh v. United States, 400 F. Supp. 238, 247 (D.D.C. 1974) (describing the standard of care for physicians); Berdyck v. Shinde, 613 N.E.2d 1014, 1017 (Ohio 1993) (describing the standard of care for nurse practitioners).
[127] *See* Darling v. Charleston Cmty. Mem'l Hosp., 211 N.E.2d 253 (Ill. 1965) (discussing direct duties of care that hospitals owe to patients); *see generally* Adelman & Robertson; 3 SUMMARY PA. JURIS. 2D *Torts*, *supra* note 121, § 37:45 (discussing theories of direct hospital liability).
[128] *Betesh*, 400 F. Supp. at 247.
[129] Heinrich v. Sweet, 308 F.3d 48, 63 (1st Cir. 2002).
[130] *See* Duckworth v. Bennett, 181 A. 558, 559 (Pa. 1935) (finding that a doctor's decision to initially diagnose a patient's complaint using physical examination instead of x-ray did not breach the standard of care when the method of diagnosis is a "matter of judgment").
[131] *Id.*

is both a cause-in-fact and legal (or "proximate") cause of the injury.[132] The provider's breach will be a cause-in-fact when "but for" the breach, the patient would not have suffered an injury.[133] Courts use different tests for proximate cause.[134] The majority of courts use the foreseeability test, which would require the patient's injury to be a foreseeable result of the healthcare provider's breach.[135]

4. Defenses

Healthcare providers using black-box AI to treat patients may assert affirmative defenses of contributory and comparative fault, which can reduce or eliminate the defendant's liability when the plaintiff's or another party's negligence caused or contributed to the plaintiff's injury.[136] Whether and how much a defendant healthcare provider's liability might be reduced by the fault of another party depends on each state's approach to contributory and comparative negligence doctrines.[137]

For example, a patient who provides incorrect information might be negligent if the provider used a black-box AI that relied on this information to produce a treatment recommendation that the provider later relied upon. Additionally, a third-party AI developer might be negligent if they provided incorrect instructions regarding the use of the black-box AI, and the provider relied on those instructions to make a recommendation. Nevertheless, because several parties may be involved with the creation and use of black-box AI, an affirmative defense of modified comparative fault raises the issue of how to apportion fault among multiple negligent defendants.

---

[132] *See* Ploch v. Hamai, 213 S.W.3d 135, 141 (Mo. Ct. App. 2006) ("To establish causation in a medical malpractice action, the plaintiff must show that the physician's conduct was the cause-in-fact and the proximate cause of the plaintiff's damages." (citing Wicklund v. Handoyo, 181 S.W.3d 143, 149 (Mo. Ct. App. 2005))).

[133] *See* JAMES UNDERWOOD, TORT LAW: PRINCIPLES IN PRACTICE 227–55 (3d ed. 2022) (discussing alternatives to the but-for test).

[134] *See id.* at 255–90 (discussing direct cause, the substantial factor, and the foreseeability tests).

[135] *Id.* at 250.

[136] *See* SPEISER ET AL., *supra* note 82, §§ 7:16, 13:2 (outlining affirmative defenses of contributory and comparative fault).

[137] *See* UNDERWOOD, *supra* note 133, at 250, 450; *Understanding Contributory Negligence and Apportionment of Fault*, BLOOMBERG L. (Feb. 13, 2023), https://perma.cc/ZW3G-2HGY (discussing the different approaches to contributory and comparative fault).

C. Proof Considerations

As a general rule, plaintiffs bear the initial burden of proof.[138] In medical malpractice cases involving black-box AI, this burden can be met in most jurisdictions using the preponderance of the evidence, "reasonable degree of medical certainty," or "reasonable degree of medical probability" standard to show that a duty was breached and caused damage.[139] However, a few jurisdictions impose a stricter interpretation of the "reasonable degree of medical certainty" standard to require proof that exceeds a probability.[140]

The majority of U.S. courts will ease the plaintiff's burden of proving the breach of a duty in cases that justify the application of the doctrine of *res ipsa loquitur*.[141] In cases involving black-box AI, *res ipsa loquitur* would allow the jury (or judge in a bench trial) to infer a breach when (1) the black-box AI that caused the injury is in exclusive control of the healthcare provider, (2) the kind of injury suffered typically does not occur in the absence of negligence, (3) the injured patient was not contributorily negligent, and (4) the healthcare provider is in a superior position to explain the accident.[142] Additionally, in cases when the defendant healthcare provider violated a statute or other regulation that was intended to protect the plaintiff against the type of injury caused by the black-box AI, the plaintiff might benefit from a rebuttable presumption of a breach under the "negligence per se" doctrine.[143]

---

[138] *See* 61 AM. JUR. 2D *Physicians, Surgeons, Etc.* § 309, Westlaw (database updated May 2023) ("The plaintiff bears the burden of proving that a physician committed medical malpractice.").

[139] *See* James v. McHenry, 828 So. 2d 94, 95 (La. Ct. App. 2002) (preponderance of evidence); Hoffman v. Carter, 648 S.E.2d 318, 326 n.9 (Va. Ct. App. 2007) (noting that "to a reasonable medical certainty" is the same as the preponderance of evidence standard); Morsicato v. Sav-On Drug Stores, Inc., 111 P.3d 1112, 1116 (Nev. 2005) (concluding that "medical expert testimony regarding standard of care and causation must be stated to a reasonable degree of medical probability").

[140] *See* Corrado v. T. Jefferson Univ. Hosp., 790 A.2d 1022, 1031 (Pa. Super. Ct. 2001) (noting that reasonable degree of medical certainty standard is not met by an expert opinion rendered in "more likely than not" terms).

[141] *See* Sides v. St. Anthony's Med. Ctr., 258 S.W.3d 811, 816–18 (Mo. 2008) (en banc) (recognizing that all but four states allow the doctrine of *res ipsa loquitur* in medical malpractice cases); LA. STAT. ANN. § 9:2794 (2023) (shifting the burden of proof in cases involving *res ipsa loquitur* in Louisiana).

[142] *See* Karyn K. Ablin, Note, *Res Ipsa Loquitur and Expert Opinion Evidence in Medical Malpractice Cases: Strange Bedfellows*, 82 VA. L. REV. 325, 330–32, 352–53 (1996).

[143] *See* Stachowski v. Est. of Radman, 95 N.E.3d 542, 544 (Ind. Ct. App. 2018) (summarizing the doctrine of negligence per se).

III. LIABILITY FRAMEWORK FOR BLACK-BOX AI IN HEALTHCARE IN THE EUROPEAN UNION

Black-box AI in healthcare poses challenges to the European Union's liability framework, many of which are similar to those in the United States. This Part first carves out general principles governing liability for AI-caused medical injuries in the European Union. It then examines recent developments in the European Union's approach to liability for AI-caused injuries. Lastly, it compares the EU and U.S. frameworks governing liability for manufacturers and individual and organizational healthcare providers.

A. General Principles Governing Liability for AI-Caused Medical Injuries in the European Union

In the European Union, liability for medical injuries is governed under a combination of EU-level law and Member States' national liability law. The European Union's institutions can enact legal acts, for example in the form of Regulations and Directives, only to the extent that they are authorized to do so under the Treaty on European Union ("TEU") and the Treaty on the Functioning of the European Union ("TFEU").[144] Regulations are directly applicable in the twenty-seven EU Member States.[145] Directives, on the other hand, must first be transposed into Member States' national law.[146] As a result, references in this Article to provisions in EU Directives (or "proposed Directives") refer more specifically to those provisions as transposed into Member States' national law. Additionally, while we cannot provide an overview of each Member State's national strict and fault-based liability laws, we provide some overarching principles that govern liability for medical injuries in the EU Member States. The following first examines manufacturer liability and individual and organizational healthcare provider liability for black-box AI in healthcare under the current EU liability framework. It then considers the burden of proof.

---

[144] Consolidated Version of the Treaty on the Functioning of the European Union art. 267, Oct. 26, 2012, 2012 O.J. (C 326) 47 [hereinafter TFEU]; *see also* Udo Bux & Mariusz Maciejewski, *Sources and Scope of European Union Law*, EUR. PARLIAMENT (Apr. 2023), https://perma.cc/WB8S-L9D8 (summarizing EU law).

[145] TFEU, *supra* note 144, art. 288(2).

[146] *Id.* art. 288(3).

1. Manufacturer Liability

Black-box AI manufacturers in the European Union can be liable under strict and/or fault-based product liability.[147] Strict product liability, governed by the Product Liability Directive ("Current PLD") as transposed into Member States' national law, imposes liability on manufacturers, including manufacturers of component parts, for injuries caused by their defective products.[148] However, the Current PLD applies to "movables" and thus, according to the majority view, would only encompass an "intangible" product like black-box AI software if it was embedded in a tangible product.[149] For such products, the Current PLD would rely on consumer expectations of safety to determine defectiveness by considering (a) the AI's presentation, (b) its reasonably foreseeable use, and (c) the moment in time when the manufacturer put the product on the market.[150] The Court of Justice of the European Union has found that for medical devices marketed for use by vulnerable patients, consumer expectations of safety are "particularly high" and that manufacturers can be liable for injuries caused by such products that are merely "potentially defective."[151] Additionally, courts in Member States might also include risk-utility considerations, particularly whether or not there was a "reasonable alternative design," when determining whether a black-box AI is defective under the Current PLD's principles.[152]

The Current PLD provides for several defenses to strict product liability claims. First, even if a black-box medical system is considered a defective product at the time of the alleged injury, the manufacturer might avoid liability by proving that the defect did not exist at the time the product was placed on

---

[147] *See* Rod Freeman et al., *Product Liability and Safety in the EU: Overview*, THOMSON REUTERS PRAC. L., https://perma.cc/AYR7-KQDS (Aug. 1, 2020) (discussing product liability for manufacturers in the European Union).

[148] Product Liability Directive (PLD), Council Directive of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products (85/374/EEC), art. 3, 1985 O.J. (L 210) 29 [hereinafter Current PLD].

[149] Christiane Wendehorst & Yannic Duller, *Safety- and Liability-Related Aspects of Software*, *in* CIVIL LIABILITY FOR ARTIFICIAL INTELLIGENCE AND SOFTWARE 185, 268–69 (Mark A. Geistfeld et al. eds., 2023); *see* Current PLD, *supra* note 148, art. 2.

[150] Current PLD, *supra* note 148, art. 6(1).

[151] Case C-503/13, Bos. Sci. Medizintechnik GmbH v. AOK Sachsen-Anhalt, ECLI:EU:C:2015:148, ¶ 39 (Mar. 5, 2015); Giorgio Risso, *Product Liability and Protection of EU Consumers: Is It Time for a Serious Reassessment?*, 15 J. PRIV. INT'L L. 210, 223 (2019).

[152] Wendehorst & Duller*, supra* note 149, at 273; *see* Current PLD, *supra* note 148, art. 2.

the market.[153] Second, the manufacturer might avoid liability if they did not sell or distribute the black-box system for an "economic purpose."[154] Third, if the defect was a result of the manufacturer's compliance with mandatory regulations applicable to the black-box system, they would not be liable under the Current PLD.[155] Fourth, according to the "development risk" or "state-of-the-art" defense, the manufacturer might escape liability for a defective black-box AI in cases where "the state of scientific and technical knowledge at the time when [they] put the product into circulation was not such as to enable the existence of the defect to be discovered."[156] Notably, while the Current PLD allows the Member States to exclude the development risk/state-of-the-art as an available defense, most have adopted this defense at least partially.[157] Finally, a component part manufacturer can escape liability by proving that the defect was caused by either the defective design of the finished black-box system or the instructions provided by the manufacturer of the finished black-box system.[158]

Fault-based product liability is governed by Member States' general fault-based liability law and is not precluded by the Current PLD.[159] The common core elements of a tort claim are (1) fault, (2) causation, and (3) damage.[160] A black-box AI manufacturer's fault will be judged objectively based on the position of a reasonable person under the circumstances.[161] If a manufacturer violates a regulation designed to protect against the risk of an AI-caused injury, this violation will weigh in favor of finding that the manufacturer's behavior was

---

[153] Current PLD, *supra* note 148, art. 7(b).
[154] *Id.* art. 7(c).
[155] *Id.* art. 7(d).
[156] *Id.* art. 7(e).
[157] *See id.* art. 15(b) (allowing Member States to exclude the development risk/state-of-the-art defense); FONDAZIONE ROSSELLI, ANALYSIS OF THE ECONOMIC IMPACT OF THE DEVELOPMENT RISK CLAUSE AS PROVIDED BY DIRECTIVE 85/374/EEC ON LIABILITY FOR DEFECTIVE PRODUCTS 38 (2004), https://perma.cc/N9DX-6LXR (reporting that all but two Member States—Finland and Luxembourg—have at least partially adopted the development risk defense).
[158] Current PLD, *supra* note 148, art. 7(f).
[159] *See* Freeman et al., *supra* note 147 (noting that Current PLD does not exclude negligent product liability under Member States' national laws).
[160] *See* Bernhard A. Koch, *Medical Liability in Europe: Comparative Analysis*, *in* MEDICAL LIABILITY IN EUROPE 611, 626–39 (Bernhard A. Koch ed., 2011).
[161] *See* Mark Geistfeld et al., *Comparative Law Study on Civil Liability for Artificial Intelligence*, *in* CIVIL LIABILITY FOR ARTIFICIAL INTELLIGENCE AND SOFTWARE 1, 46–48 (Mark A. Geistfeld et al. eds., 2023) (noting that fault is determined using an objective standard).

unreasonable under the circumstances. As a result, black-box AI manufacturers will likely be liable if they were unreasonable in their design, manufacture, instructions, warning, and monitoring as they relate to the AI that caused an injury.

2. Healthcare Provider Liability

Tort liability for medical injuries caused by healthcare providers is governed by fault-based principles in all EU Member States.[162] As a result, when the fault of an individual healthcare provider causes patient damage, the provider will be liable under Member States' fault-based tort law. Organizational healthcare providers like hospitals can be vicariously liable for the fault of the individual healthcare providers in their employment, but they can also be directly liable for faults related to personnel, equipment, or facilities.[163]

Individual healthcare providers that use black-box AI to treat patients are at fault if they fail to fulfill the objective standard of care in their corresponding field or *lex artis*.[164] This standard of care takes into account the state of medical science at the time of the treatment.[165] As a result, the standard of care for individual healthcare providers that use black-box AI might consider the providers' limited ability to predict and understand the AI's output.[166] Healthcare organizations that use black-box AI might also be directly liable for injuries that stem from organizational deficiencies.[167] Finally, whether an individual or organizational healthcare provider is at fault in a given case involving a black-box AI will likely be impacted by legislative safety provisions[168] and medical professional rules or guidelines.[169]

---

[162] *See id.* at 44 ("All European legal systems recognize (tortious) fault-based liability."). *But see* Koch, *supra* note 160, at 531–33 (noting that while Sweden recognizes fault-based medical tort liability, most medical malpractice cases are handled under a no-fault insurance compensation scheme).

[163] *See* Koch, *supra* note 160, at 630 (noting that in addition to vicarious liability, hospitals can also be directly liable for organizational fault).

[164] *Id.* at 628.

[165] *Id.* at 629.

[166] *See* Geistfeld et al., *supra* note 161, at 56–57.

[167] *See* Koch, *supra* note 160, at 630 (discussing hospital or private practice liability); DIETER GIESEN, INTERNATIONAL MEDICAL MALPRACTICE LAW: A COMPARATIVE LAW STUDY OF CIVIL LIABILITY ARISING FROM MEDICAL CARE 58–59 (1988) (discussing direct hospital liability).

[168] Geistfeld et al., *supra* note 161, at 55.

[169] Koch, *supra* note 160, at 625.

A medical injury involving an individual or organizational healthcare provider's use of a black-box AI must be caused by the provider's fault for liability to attach.[170] Factual causation is satisfied when a healthcare provider's breach of the standard of care is a "*condicio sine qua non*" (or but-for cause) of the patient's injury.[171] However, even if the but-for test cannot clearly establish causation in a case involving a black-box AI, some Member States may still impose liability if the healthcare provider's fault increases the risk of the occurrence of injury or results in a loss of a chance for the patient to avoid the injury that occurred.[172] On the other hand, a healthcare provider's breach of the standard of care might not be considered a legal cause of the patient's injury if it does not bear a sufficient level of "blameworthiness."[173] Generally, the provider will be liable if the patient's injury is considered a "natural consequence" of their breach.[174] In some EU Member States, this will be the case even if the injury is considered to be "a highly improbable consequence," while others might exclude liability for injuries that are "extraordinarily unlikely" from the perspective of a healthcare provider with relevant expertise immediately prior to the breach in question.[175]

---

[170] Geistfeld et al., *supra* note 161, at 26.

[171] GIESEN, *supra* note 167, at 177; *see* Herman Nys, *Medical Liability in Belgium*, *in* MEDICAL LIABILITY IN EUROPE 61, 73–74 (Koch ed., 2011) (describing Belgium's "theory of equivalence"); BASIL S. MARKESINIS ET AL., THE GERMAN LAW OF TORTS: A COMPARATIVE TREATISE 61 (5th ed. 2019) (describing Germany's method of elimination).

[172] *See, e.g.*, Attila Menyhárd, *Medical Liability in Hungary, in* MEDICAL LIABILITY IN EUROPE 291, 301 (Bernhard A. Koch ed., 2011) (noting Hungary's approach to causation and liability for increasing the risk of injury); Ivo Giesen & Esther Engelhard, *Medical Liability in the Netherlands, in* MEDICAL LIABILITY IN EUROPE 361, 376–77 (Bernhard A. Koch ed., 2011) (noting that the burden of proof for causation shifts in the Netherlands in cases of increased risks of damage); Nys, *supra* note 171, at 74 (noting that Belgium provides liability for some causes in which cause-in-fact cannot be established by recognizing a claim for loss of chance); Koch, *supra* note 160, at 636 (noting that some jurisdictions in Europe eliminate causation issues by recognizing loss of chance as a "distinct damage").

[173] S*ee* Koch, *supra* note 160, at 628 (noting that causation considers both the factual link between the breach to the damage as well as the "blameworthiness" of the conduct); Miguel Martín-Casals & Josep Solé, *Medical Liability in Spain, in* MEDICAL LIABILITY IN EUROPE 451, 468–69 (Bernhard A. Koch ed., 2011) (noting that legal causation in Spain considers whether the injury is a "natural, adequate, and sufficient consequence" of the breach).

[174] *See* Giesen & Engelhard, *supra* note 172, at 375 (noting that legal causation in the Netherlands considers nature of liability and type of harm).

[175] *Compare* Giesen & Engelhard, *supra* note 172, at 379–80 (observing that a Dutch court found legal causation even though damage was "a highly improbably consequence"), *with* Martín-Casals & Solé, *supra* note 173, at 469 (explaining that there is no legal causation in Spain when injury is "extraordinarily unlikely").

In medical malpractice cases involving black-box AI, there may be multiple causes of a patient's injury. In the majority of EU Member States, when more than one individual or organizational healthcare provider has breached the standard, and those breaches are potential alternative causes of the patient's injury, the providers will be jointly and severally liable.[176] However, in cases where the black-box AI (or some other circumstance like a pre-existing condition) rather than a second healthcare provider is deemed one potential cause of the patient's injury, the healthcare provider may (1) completely escape liability or be liable for the entire damage under an "all-or-nothing" rule; or (2) be apportioned a percentage of the damage under a "proportional liability" rule.[177]

3. Proof Considerations

Under the default position, the patient will bear the burden of proof in cases involving injuries caused by black-box AI.[178] The standard of proof in many, but not all, Member States requires more than a preponderance of the evidence.[179] However, there are several situations in which the plaintiff's burden of proof is either relaxed or shifts completely to the defendant.[180] For example, the burden of proof on fault or causation might shift to the defendant if the black-box AI that caused the injury is deemed to be in the defendant's sphere of control, if the defendant is in a better position than the patient to bear the burden of proof, or in the case of "gross errors."[181]

---

[176] Koch, *supra* note 160, at 635.

[177] *See id.* (discussing the different approaches to causation in cases where at least one provider breached the standard of care, but another cause contributed to the damage).

[178] *See id.* at 630 (noting that, in general, the plaintiff bears the burden of proof). *But see* Menyhárd, *supra* note 172, at 300 (explaining that once a plaintiff proves damages and causation in Hungary, fault is presumed).

[179] *See* Koch, *supra* note 160, at 632 (noting that Italy and Sweden require probability; Austria and Poland require high probability; and Belgium, France, Czech Republic, and Spain sometimes require close to certainty).

[180] *Id.* at 633–34 (discussing presumption and burden shifting mechanisms in several EU Member States).

[181] *See* Menyhárd, *supra* note 172, at 337 (discussing evidentiary burden shifting in Hungary); Bürgerliches Gesetzbuch [BGB] [Civil Code], § 630h, https://perma.cc/Y336-HAXE (Ger.) (listing provisions for shifting the evidentiary burden in medical malpractice cases in Germany).

B. Recent Developments in the European Union's Approach to Liability for AI-Caused Injuries

With regard to liability for AI-caused injuries, the European Union has recognized that an AI's autonomous behavior, data dependency, opacity, lack of interpretability, product and system complexity, continuous adaptation and software updates, and lack of predictability, present liability challenges.[182] As a result, it can be difficult to connect potentially problematic AI decisions made to a legally responsible party, making it difficult for injured persons to obtain compensation under the current liability framework.[183] The Proposed PLD and Proposed AILD represent the EC's recent attempt to present a more harmonized solution to these liability challenges. An overarching goal of the two newly proposed Directives is to govern risks presented in the digital age in a manner that balances the interests of beneficial innovation with the rights of injured parties.[184] The Proposed PLD attempts to do this by "ensuring a fair balance between the legitimate interests of manufacturers, injured persons and consumers in general."[185] The Proposed AILD seeks to do this by simultaneously ensuring compensation for AI-caused damage and providing more clarity surrounding liability for risks associated with AI.[186] The following first analyzes product liability under the Proposed PLD and then fault-based liability under the Proposed AILD.

---

[182] *See Report from the Commission to the European Parliament, the Council and the Economic and Social Committee on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics*, COM (2020) 64 final (Feb. 19, 2020) [hereinafter Report from the Commission]; *Commission Staff Working Document Impact Assessment Report Accompanying the Document Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence*, COM (2022) 496 final (Sept. 28, 2022) [hereinafter Commission Staff Working Document].

[183] *Commission White Paper on Artificial Intelligence - A European Approach to Excellence and Trust*, COM (2020) 65 final (Feb. 19, 2020).

[184] *See Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Explanatory Memorandum,* COM (2021) 206 final (Apr. 21, 2021) [hereinafter Proposed AI Act]; Proposed AILD, *supra* note 15, Explanatory Memorandum.

[185] Proposed PLD, *supra* note 14, Explanatory Memorandum § 1.1.

[186] *See* Proposed AILD, *supra* note 15, Explanatory Memorandum (noting the aim of reducing uncertainty regarding liability for developers of AI).

1. Product Liability Under the Proposed Product Liability Directive

The Proposed PLD retains the main structure for EU-level product liability by providing for strict liability when defective products cause damage. Member States' national laws governing strict product liability cannot be more or less restrictive than those set forth in the Proposed PLD.[187] The Proposed PLD explicitly includes software in its "product" definition.[188] Under the Proposed PLD, a product is defective if it "does not provide the safety which the public at large is entitled to expect, taking all circumstances into account."[189] While similar to the Current PLD's "consumer expectations" test, the Proposed PLD adds new factors that might affect defectiveness, including, among others, (a) the product's presentation and instructions, (b) the product's "reasonably foreseeable use and misuse," (c) the product's ability to continually learn, (d) the product's effect on other products with which it is reasonably expected to be used, (e) the time when the manufacturer places the product on the market or loses control over the product, (f) "product safety requirements," (g) regulatory intervention related to the product's safety, and (h) expectations of the intended end-users, who are often consumers.[190] The term "manufacturer's control" includes post-market software updates, upgrades, and modifications.[191]

New evidentiary rules in the Proposed PLD will make it more difficult for an AI manufacturer to avoid liability. First, the Proposed PLD provides a mechanism for injured parties with plausible claims to obtain evidence from defendants, thus alleviating the information asymmetry that victims may encounter in cases involving black-box AI.[192] Second, the Proposed PLD provides for a presumption of defectiveness and/or causation when the claimant faces "excessive difficulties" in proving one or both of these elements "due to technical or scientific complexity" of a product that likely contained a damage-causing defect.[193] A patient injured by a black-box AI would be entitled

---

[187] Proposed PLD, *supra* note 14, art. 3.
[188] *Id.* art. 4(1).
[189] *Id.* art. 6(1).
[190] *Id.*
[191] *Id.* art. 4(5).
[192] *Id.* art. 8(1). However, when the disclosure of such evidence would reveal trade secrets, the court ordering the disclosure can offer the requisite confidentiality protection. *Id.* art. 8(4).
[193] *Id.* art. 9(4).

to a presumption of defectiveness if the manufacturer failed to either disclose required evidence, failed to comply with a safety standard intended to protect against the patient's injury, or if the patient's injury was caused by an "obvious malfunction" in the normal course of the AI's use.[194] Additionally, a presumption of causation would apply if the injury constitutes damage "of a kind typically consistent" with the defect.[195]

A manufacturer may not be liable under the Proposed PLD if they "did not place the product on the market or put it into service."[196] Additionally, a manufacturer may not be liable if it is probable that the product was not defective when they placed it on the market or when the manufacturer could not have discovered the defect considering the "objective state of scientific and technical knowledge" at that time.[197] However, a party that subsequently substantially modifies a product will be considered a "manufacturer" and will be strictly liable for injuries caused by the modification.[198]

2. Fault-Based Liability Under the Proposed AI Liability Directive

The Proposed AILD provides fault-based liability rules for AI-caused damages that are unrelated to defective products.[199] Unlike the Proposed PLD, the Proposed AILD does not prohibit Member States from imposing fault-based liability laws that are stricter than those contained in the newly proposed Directive.[200] The Proposed AILD relies on the substantive duties of care already established by the EU Member States' national law. A determination of fault is still based on non-compliance with an existing "duty of care under Union or national law."[201] The Proposed AILD defines "duty of care" as "a required

---

[194] *See id.* art. 9(2) (discussing conditions where defectiveness of a product shall be presumed).

[195] *See id.* art. 9(3) ("The causal link between the defectiveness of the product and the damage shall be presumed, where it has been established that the product is defective and the damage caused is of a kind typically consistent with the defect in question.").

[196] *Id.* art. 10(1)(a).

[197] *Id.* art. 10(1)(c), (e).

[198] *Id.* arts. 7(4), 10(1)(g).

[199] Proposed AILD, *supra* note 15, Explanatory Memorandum.

[200] *Id.* art. 1(4); *id.* recital 14.

[201] *Id.* Explanatory Memorandum; *see id.* recital 23 ("Such a fault can be established in

standard of conduct, set by national or Union law, in order to avoid damage to legal interests recognized at national or Union law level, including life, physical integrity, property and the protection of fundamental rights."[202] Finally, injuries caused by a breach of a duty of care that occurs subsequent to an AI's output appear to fall outside of the Proposed AILD's scope.[203]

The Proposed AILD seeks to decrease the evidentiary burden on plaintiffs by enabling Member State courts to order defendants and third parties to produce evidence that the court determines is "necessary and proportionate" to a claim for damage caused by a high-risk AI system.[204] The Proposed AILD defines a "high-risk AI system" by referring to the Proposed AI Act.[205] Under the Proposed AI Act, black-box medical AI will qualify as a high-risk system because it will require "third-party conformity assessment" prior to market authorization[206] as either a class IIa, IIb, or III "medical device" under the Medical Device Regulation ("MDR")[207] or as a class B, C or D "in vitro diagnostic

---

respect of non-compliance with Union rules which specifically regulate high-risk AI systems like the requirements introduced for certain high-risk AI systems by [the AI Act], requirements which may be introduced by future sectoral legislation for other high-risk AI systems according to [Article 2(2) of the AI Act], or duties of care which are linked to certain activities and which are applicable irrespective whether AI is used for that activity. At the same time, this Directive neither creates nor harmonises the requirements or the liability of entities whose activity is regulated under those legal acts, and therefore does not create new liability claims. Establishing a breach of such a requirement that amounts to fault will be done according to the provisions of those applicable rules of Union Law, since this Directive neither introduces new requirements nor affects existing requirements." (alterations in original)).

[202] *Id.* art. 2(9).

[203] *See id.* recital 15 ("[T]his Directive should only cover claims for damages when the damage is caused by an output or the failure to produce an output by an AI system through the fault of a person, for example the provider or the user under [the AI Act]. There is no need to cover liability claims when the damage is caused by a human assessment followed by a human act or omission, while the AI system only provided information or advice which was taken into account by the relevant human actor. In the latter case, it is possible to trace back the damage to a human act or omission, as the AI system output is not interposed between the human act or omission and the damage, and thereby establishing causality is not more difficult than in situations where an AI system is not involved." (second alteration in original)); Mindy Nunez Duffourc & Sara Gerke, *The Proposed EU Directives for AI Liability Leave Worrying Gaps Likely to Impact Medical AI*, 6 NPJ DIGIT. MED., Apr. 26, 2023, at 3 fig.2 (noting that Proposed AILD recital 15 appears to exclude claims based on a breach of a duty of care that occurred after the AI's output from the Directive's scope).

[204] Proposed AILD, *supra* note 15, art. 3(4).

[205] *Id.* art. 2(2).

[206] Proposed AI Act, *supra* note 184, art. 6(1)–(2); *id.* annex II(11)–(12).

[207] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on Medical Devices, Amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and Repealing Council Directives 90/385/EEC and 93/42/EEC, art. 2(1), 2(4), annex VIII 2.5, 3.5, 3.7, 6.2-3, 2017 O.J. (L 117) 1 [hereinafter MDR].

medical device" under the In Vitro Diagnostic Medical Device Regulation ("IVDR").[208] Failure to comply with an order to produce evidence can result in a rebuttable presumption of fault for the party on the matter which the evidence was intended to prove.[209]

Finally, the Proposed AILD provides for a rebuttable presumption that a defendant's fault caused the AI's injury-causing output or lack thereof.[210] This presumption applies when (1) the fault is "reasonably likely" to have "influenced" the AI's output, and (2) the "output gave rise to the damage."[211] Because fault must influence the AI's output, the Proposed AILD does not apply, and thus does not provide a presumption of causation, when human fault follows, rather than precedes, the AI output because in that case, the "output is not interposed between the human act or omission and the damage."[212]

C. Comparative Analysis of the EU and U.S. Liability Frameworks

Despite the various approaches in EU Member States, there is some concordance on the basic liability rules that govern medical injuries between both the EU Member States themselves and the approach in the United States. The two tables below compare the current U.S. and EU frameworks governing liability for black-box AI manufacturers (Table 1) and individual and organizational healthcare providers using black-box AI (Table 2). The yellow highlights in both tables indicate areas of law in EU Member States that will be affected by the Proposed PLD and Proposed AILD.

---

[208] Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in Vitro Diagnostic Medical Devices and Repealing Directive 98/79/EC and Commission Decision 2010/227/EU, art. 2(2), annex VIII, 2017 O.J. (L 117) 176 [hereinafter IVDR].
[209] Proposed AILD, *supra* note 15, art. 3(5).
[210] *Id.* art. 4(1).
[211] *Id.* art. 4(1)(b), (c).
[212] *See id.* art. 4(1); *id.* recital 15.

Tort Liability for Manufacturers	**United States**		**European Union**	
Type of Liability	**Strict Product**	**Fault-Based**	**Strict Product**	**Fault-Based**
Level of Substantive Governing Law	Primarily State	Primarily State	Primarily EU (through Proposed PLD as transposed into Member State Law)	Member State *Proposed AILD introduces new proof rules to be transposed in Member State law.
Elements of Liability	Product defect, causation, and damages	Manufacturer negligence, causation, and damages	Product defect, causation, and damages	Manufacturer negligence, causation, and damages
Types of Product Defects	Design, manufacturing, and marketing	N/A	Design, manufacturing, and marketing	N/A
Types of Manufacturer Negligence	N/A	Negligent design, manufacturing, and marketing	N/A	Negligent design, manufacturing, and marketing
Legal Standard for Determining Product Design Defects	Consumer expectations and/or risk-utility	N/A	Consumer expectations and/or risk-utility *Proposed PLD adds new factors that can impact whether a product is defective.	N/A
Legal Standard for Determining Whether a Manufacturer Is Negligent	N/A	Breach of a duty of care	N/A	Breach of a duty of care
Liability for Manufacturers of Component Parts	Defective design, manufacturing, marketing of component parts. No *design defect* liability if designed to manufacturer's specification and had no actual or constructive	Negligent design, manufacturing, and marketing of component parts	Defective design, manufacturing, marketing of component parts. No liability if defect was caused by *design or*	Negligent design, manufacturing, and marketing of component parts

	knowledge of design defect. No *manufacturing defect* liability if component part underwent substantial change when incorporated into final product. No *marketing defect* liability for failing to warn end user if no actual or constructive knowledge that component part would make product unreasonably dangerous.		*instructions* of final product.	
Defenses	• Preemption • Unavoidably unsafe product • State-of-the-art • Learned Intermediary Doctrine • Consumer misuse, alteration, or modification • Consumer negligence • Assumption of risk	• Preemption • Unavoidably unsafe product • State-of-the-art • Learned Intermediary Doctrine • Consumer misuse, alteration, or modification • Consumer negligence • Assumption of risk	• No defect on market entry *Proposed PLD also considers whether the manufacturer has control over the product even after market entry. • No intention to sell or distribute for economic purpose • Defect caused by compliance with mandatory regulations • Development risk/state-of-the-art	Varies with Member State law

Table 1: Tort Liability for Black-Box AI Manufacturers
in the United States and European Union

Medical Liability for Healthcare Providers	**United States**	**European Union**
Type of Liability	Fault-based (primarily tort)	Fault-based (primarily contract)
Level of Substantive Governing Law	State	Member State *Proposed AILD introduces new proof rules to be transposed in Member State law.
Elements of Liability	Duty, breach, causation, damage	Fault, causation, damage
Standard of Proof	Probability	Generally more than probability
Burden of Proof	Plaintiff	Plaintiff
Deviations from General Burden of Proof Rules	Relaxation of plaintiff's burden of proof on negligence under *res ipsa loquitur* and negligence *per se* doctrines.	Relaxation of plaintiff's burden of proof on fault and causation in various circumstances. *Proposed AILD provides AI-specific burden shifting mechanisms.
Legal Standard for Determining Whether a Healthcare Provider is at Fault	The level of care exercised by medical providers in the same field.	The level of care exercised by medical providers in the same field.
Types of Healthcare Providers	Individual, organizational	Individual, organizational

Table 2: Medical Liability for Individual and Organizational Healthcare Providers Using Black-Box AI in the United States and European Union

IV. LIABILITY ISSUES FOR BLACK-BOX AI IN THE UNITED STATES WITH INSIGHTS FROM THE EUROPEAN UNION

To demonstrate the applicability and limitations of the current U.S. liability framework in medical injury cases that involve black-box AI, this Part analyzes three hypothetical scenarios under U.S. tort liability law and the current and newly proposed EU liability framework to determine who (if anyone) may be

held liable for a person's injury. Scenario 1 focuses on an *autonomous* black-box AI (so its output is *not* reviewed or overseen by a human) that is used by a healthcare provider for patient treatment. Scenario 2 involves a *non-autonomous* black-box AI used by a healthcare provider, and Scenario 3 deals with a *DTC* black-box AI.[213]

A. Scenario 1: Autonomous Black-Box AI Used by a Healthcare Provider for Patient Treatment

Imagine a legally-marketed autonomous black-box AI that uses deep learning to read head CT scans without the input of a radiologist is used by an individual or organizational healthcare provider to treat patients.[214] If the AI reports a positive finding, the patient is referred to a specialist for further evaluation and treatment. However, negative CT scans are not reviewed further. In this Scenario, the black-box AI incorrectly reports a negative CT scan with no significant findings, which leads to a missed diagnosis and subsequent injury for the patient.

1. Analysis Under Current U.S. Tort Liability Law

In this Scenario, the patient could potentially sue the manufacturer under product liability law and/or the individual or organizational healthcare providers under fault-based negligence law.

a. Manufacturer Liability

To recover against the manufacturer under either strict or negligent product liability law, the patient would need to prove that this autonomous black-box AI suffered from a design, manufacturing, or marketing defect that

---

[213] For more information on the different terminology (autonomous, non-autonomous, DTC), see *supra* Sections I.A, I.B.2.

[214] *See* U.S. FOOD & DRUG ADMIN., *supra* note 42 (describing autonomous AI that diagnoses diabetic retinopathy); Nicole Wetsman, *Autonomous X-Ray-Analyzing AI Is Cleared in the EU*, VERGE (Apr. 5, 2022), https://perma.cc/2VJN-Q4HD (describing autonomous AI that reads chest X-rays); Shammi Ramlakhan et al., *Understanding and Interpreting Artificial Intelligence, Machine Learning and Deep Learning in Emergency Medicine*, 39 EMERGENCY MED. J. 380, 380–84 (2022) (discussing potential of AI to recognize fractures in X-rays).

caused their injury.[215] We assume for this Section that the AI system would be considered a "product" subject to strict liability law.

A *design defect* claim would require the patient to prove that the AI's design made it unreasonably dangerous.[216] This will be difficult if the manufacturer designed this AI specifically to engage in noninterpretable and potentially unpredictable algorithmic reasoning to ultimately produce CT reports. This AI likely responded to vast amounts of data during the training process and, to some extent, *designed itself* using deep learning. Nevertheless, the manufacturer still had some control over the design of the AI's source algorithms, training, and validation and, thus, could be liable for design defects if these processes caused the AI to misread the CT. For example, if the manufacturer designed the AI to use poor-quality training data, it was likely foreseeable that the AI would produce poor-quality CT reports in accordance with the "garbage in, garbage out" adage.[217] On the other hand, if this AI's uninterpretable and potentially unpredictable reasoning is a feature of its design, this unpredictability might factor into its usefulness, and the cost of eliminating this unpredictability may have a detrimental effect on the AI's accuracy, tilting risk-utility considerations against finding a design defect. Even assuming a design defect, the plaintiff would have to show that a reasonable alternative design existed at the time the AI was placed on the market. This would require the plaintiff to obtain and understand information surrounding the complex technical aspects of AI architecture, an evidentiary burden that would be extremely difficult to carry.[218]

Liability for a *manufacturing defect* would require the plaintiff to prove that the black-box AI system was not operating as designed when it misread the CT scan.[219] For example, if the AI was designed to be trained with data from a diverse patient population but was instead trained with data from a homogenous patient population, this might be considered a manufacturing defect. The potentially "defective" data set might also be considered a component part if it was provided by a third party. Of course, if a physical

---

[215] *See supra* Section II.A. For negligent product liability, an additional requirement that the manufacturer's negligence caused the defect would apply.
[216] *See supra* Section II.A.1.
[217] *See* Pujol, *supra* note 19, at 31–34 (discussing the "garbage in, garbage out" adage in machine learning).
[218] *See supra* Section II.C.
[219] *See supra* Section II.A.2.

hardware component caused the AI to produce an incorrect CT report, for example, a faulty connection component that disabled a crucial feature of the AI's algorithm, this might also be a manufacturing defect. However, if the AI was operating as originally designed by the manufacturer, it would likely not suffer from a manufacturing defect.

A *failure to warn* claim against the manufacturer would require the patient's injuries to be foreseeable.[220] While the manufacturer might be able to foresee that this AI could miss a positive finding on a CT scan generally, it may not be able to predict the specific risks that caused the AI to miss the CT finding in this patient's case. However, if the manufacturer was aware or became aware of the specific risks that caused the AI to produce this patient's inaccurate CT report, even after the AI was on the market, it would likely be liable for failing to warn of such risks. On the other hand, if the manufacturer issued warnings about the general risk that the AI may not produce an accurate CT report in every single case, it will likely not be liable for failing to warn of the more specific risk that manifested in this individual case if such a risk was unforeseeable because the manufacturer could not predict how the AI would respond in each individual case.

Even if the AI was defective, the manufacturer may have *several defenses*.[221] First, if this black-box AI received PMA from the FDA, product liability claims would likely be preempted. Second, even if this AI took a less stringent premarket review path and received only a 510(k) clearance, the manufacturer might still avoid product liability if its ability to autonomously read CT scan rendered it an unavoidably unsafe product that, despite "some risks, perhaps serious ones," has the potential to "save lives and reduce pain and suffering."[222] Third, if at the time the AI was manufactured, it complied with applicable safety regulations, and it was not technologically possible to design a less risky and comparably beneficial AI to read CT reports, the manufacturer may avoid liability under the state-of-the-art defense. Fourth, if the manufacturer provided sufficient warnings to the individual or organizational healthcare provider that deployed the AI to read the patient's CT scan, the

---

[220] *See supra* Section II.A.3.
[221] *See supra* Section II.A.4.
[222] *See* Brown v. Super. Ct. of S.F., 751 P.2d 470, 479 (Cal. 1988) (discussing unavoidably unsafe drugs).

provider's decision to use the AI in the patient's treatment may shield the manufacturer from liability under the LID. Finally, a host of subsequent actions by other parties that constitute misuse, alteration, modification, or negligence, might foreclose product liability for the manufacturer. This might include such things as the radiology technician's failure to properly position the patient for the CT or the provider's decision to use the CT scan in this patient's treatment in contradiction to the manufacturer's instructions.

b. Healthcare Provider Liability

A negligence claim based on the autonomous black-box AI's missed finding would also face several challenges. First, the AI is performing actions that, if performed by a human radiologist, would be considered medical treatment—in this Scenario, reading CT scans and producing radiology reports.[223] Currently, AI is not capable of negligence because, unlike healthcare providers, they are not legal persons.[224]

Although this AI is autonomous, the individual or organizational healthcare provider(s) that chose to use the AI might still be liable for injuries caused by the AI's failure to report positive CT findings if they were negligent in using and/or implementing the AI into clinical care generally or in this patient's treatment specifically.[225] On the other hand, if neither provider's decision to use the AI nor their implementation of the AI in clinical practice was negligent, they may not be held liable for an injury caused by the AI's inaccurate CT report.[226] Even if the healthcare provider was negligent, for example, by using the AI in the wrong patient population, failing to train staff regarding the proper

---

[223] *See* Mindy Nunez Duffourc & Dominick S. Giovanniello, *The Autonomous AI Physician: Medical Ethics and Legal Liability*, *in* MULTIDISCIPLINARY PERSPECTIVES ON ARTIFICIAL INTELLIGENCE AND THE LAW 207 (Henrique Sousa Antunes et al. eds., 2024) (noting that individualized medical consultation, diagnosis, and prognosis are "tasks that fall squarely within the scope of medical practice").

[224] *See* Jason Chung & Amanda Zink, *Hey Watson—Can I Sue You for Malpractice? Examining the Liability of Artificial Intelligence in Medicine*, ASIA PAC. J. HEALTH L. & ETHICS, Mar. 2018, at 51, 67 (2018) ("Courts have traditionally deemed it impossible for machines to have legal liability as they are not legal persons. . . . Regarding the human-centric categories, both negligence and vicarious liability as causes of action appear to require personhood.").

[225] *See supra* Section II.B.2.

[226] *See* Duffourc, *supra* note 17, at 3–4, 17 (noting that injury caused by an autonomous medical AI might not be compensated under existing product liability or medical negligence causes of action in U.S. tort law when there is no healthcare provider negligence); *see generally* Duffourc & Gerke, *supra* note 203 (finding a similar liability gap in EU law).

use and oversight of the system, or failing to update the system, it could still be difficult to prove that this negligence caused the AI's inaccurate CT report. Establishing a causal link between the provider's negligence and the AI's subsequent CT report might require the plaintiff to understand and evaluate highly technical aspects of the AI's deep learning architecture.[227] Finally, the liability of a negligent provider in this Scenario could be reduced or barred by the negligence of other parties, including the patient.[228] This might be the case, for example, if the patient was advised to follow up for further medical treatment and failed to do so and such failure caused or contributed to their injury remaining undiagnosed or untreated.

2. Analysis Under the Current and Newly Proposed EU Liability Framework

In this Scenario, the patient could potentially sue the manufacturer for product liability and/or the individual or organizational healthcare providers for fault-based medical liability under Member States' national law, which incorporates the Current PLD and would incorporate the Proposed PLD and Proposed AILD once they are enacted and transposed into national law.

a. Manufacturer Liability

This autonomous black-box AI also challenges the European Union's current product liability framework for manufacturers. Under the Current PLD, if this AI is considered intangible software and is not embedded into a tangible product, it would not be considered a product.[229] Assuming that this AI is a product, it might not be defective under the current consumer expectations and risk-utility tests.[230] For example, if, at the time it was placed on the market, there was no reasonably alternative design that would have offered comparable benefits with less risk, the AI might not be considered defective. As in the United States, this would present a significant barrier for the patient's

---

[227] *See supra* Sections II.B.3, II.C.
[228] *See supra* Section II.B.4.
[229] *See* Current PLD, *supra* note 148, art. 2 (discussing the scope of the Directive); *see also supra* Section III.A.1.
[230] *See supra* Section III.A.1.

product liability claim because it requires them to understand and obtain highly technical and potentially vast information about the AI's design and operation to overcome their burden of proving that the AI was defective. Again, as in the United States, even if this AI was deemed defective under the Current PLD, the manufacturer might be able to avoid liability under the state-of-the-art defense.[231]

The patient has a better product liability claim under the Proposed PLD. First, the Proposed PLD explicitly includes software in its product definition.[232] Second, the Proposed PLD could shift some of the evidentiary burden associated with the complex technical aspects of the AI's architecture to the manufacturer, making it easier for the patient to obtain information and prove that the AI's missed CT finding was the result of a defect.[233] Third, the Proposed PLD's recognition that both the AI's design and the relationship between the AI and its manufacturer are dynamic might help the patient prove a defect. For example, if this AI is not locked and has the ability to continuously learn, this factor might weigh in favor of finding it defective.[234] Additionally, if the manufacturer failed to provide a necessary post-market update that would have prevented the AI from missing the positive CT finding, the AI might be considered defective under the Proposed PLD, which recognizes the potential for the manufacturer to retain control over the AI after market entry. Finally, this AI could also be deemed defective under the Proposed PLD if the manufacturer failed to comply with mandatory safety requirements, such as those set forth in the European Union's Proposed AI Act and current Medical Device Regulation.

---

[231] *See* Current PLD, *supra* note 148, art. 7(e) (stating that economic operators will not be liable for damages caused by a defective product if they prove "in the case of a manufacturer, that the objective state of scientific and technical knowledge, at the time the product was placed on the market, put into service, or within the manufacturer's control, was not such that the defectiveness could not be discovered"); *see supra* Section III.A.1.

[232] *See* Proposed PLD, *supra* note 14, art. 4(1) (noting "'[p]roduct' includes electricity, digital manufacturing files and software"); *see also supra* Section III.B.1.

[233] *See* Proposed PLD, *supra* note 14, art. 8(1) (stating that upon the request of the injured person who has presented "evidence sufficient to support the plausibility" of their compensation claim, the courts may order the defendant to disclose relevant evidence at its disposal); *see also supra* Section III.B.1.

[234] *See* Proposed PLD, *supra* note 14, art. 6(1) (stating that a product is "defective when it does not provide the safety which the public at large is entitled to expect"); *see also supra* Section III.B.1.

Similar to U.S. law, the manufacturer could, however, still escape liability even under the Proposed PLD if the black-box nature of the AI's reasoning process made it impossible for the manufacturer to predict that AI would misread the CT scan. For example, the fact that this AI uses black-box technology to make decisions that the manufacturer may not be able to reasonably predict or prevent signals a lack of manufacturer control and weighs against finding a defect.[235] Additionally, if the manufacturer fully complied with the applicable safety standards and the patient was appropriately informed of both medical and technical risks associated with using the AI, the missed CT finding may not be considered a product defect.[236] Finally, the Proposed PLD retains the state-of-the-art defense, which might allow the manufacturer to escape liability if the AI's missed CT finding resulted from a defect that the manufacturer could not have discovered when the AI was placed on the market.

A product liability claim based on negligence under either the current national fault-based liability law or national fault-based liability law with the Proposed AILD would encounter many of the same challenges related to the product's design that make strict liability claims under the Proposed PLD challenging. Namely, it might be difficult to prove negligence in the product's design if the AI was designed to be uninterpretable and unpredictable and its reasoning process was unforeseeable. However, if the manufacturer was negligent in selecting the AI's training data, for example, then the Proposed AILD would entitle the patient to a presumption that the poor training data caused the AI's inaccurate CT report for this patient as long as it is reasonably likely that the negligence caused the AI's failure.[237]

b. Healthcare Provider Liability

Individual and organizational healthcare providers might also escape fault-based liability in the European Union despite the Proposed AILD's goal of facilitating fault-based liability claims in cases involving an AI-caused injury. This is because the Proposed AILD would not alter the basic elements of proving

---

[235] *See supra* note 190 and accompanying text.

[236] *See supra* note 190 and accompanying text (noting that new factors for determining defectiveness in the Proposed PLD art. 6(1) include compliance with safety standards and consumer expectations).

[237] *See* Proposed AILD, *supra* note 15, art. 4(1); *see also supra* Section III.B.2.

fault under current national law, which for most Member States, as in the United States, would require the patient to first prove that an individual or organizational healthcare provider failed to comply with an existing duty of care.[238] Since the AI itself is not considered a legal person, the patient must connect the AI's missed CT finding to the healthcare provider(s) that chose to use the AI in the patient's care. For example, if the healthcare provider used the AI in the wrong patient population, there would likely be a breach of the standard of care. However, if the healthcare provider's use, implementation, and oversight of the AI complied with its existing duties of care, for example, by following the manufacturer's instructions, properly training staff that use the AI, and monitoring the AI's performance, then the healthcare provider may also not be at fault.[239]

If the patient is able to establish fault by an individual or organizational healthcare provider, the Proposed AILD's evidentiary presumptions on causation would benefit the patient.[240] Considering the above example of a breach stemming from the provider's use of the AI in the wrong patient population, the patient would benefit from a presumption that this breach caused the AI to miss a significant finding in the patient's CT scan as long as the missed finding was a "reasonably likely" result of using the AI in the wrong patient population.[241] This presumption of causation will help reduce the barriers associated with obtaining and understanding evidence related to the complex technical inner-workings of the AI at the causation stage.

B. Scenario 2: Non-Autonomous Black-Box AI Used by a Healthcare Provider

Imagine a legally-marketed non-autonomous black-box AI that quickly predicts the cause of acute respiratory failure in a patient presenting to the

---

[238] *See* Proposed AILD, *supra* note 15, Explanatory Memorandum; *id.* recitals 23–24; *see also supra* Section III.B.2.
[239] See *supra* notes 201–02 and accompanying text.
[240] *See* Proposed AILD, *supra* note 15, art. 4(1) (stating that "national courts shall presume . . . the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output," where certain conditions are met).
[241] *See id.*

emergency room, a task that can be difficult for human physicians.[242] A human physician then relies upon the AI's prediction to determine the appropriate emergency medical treatment. The AI predicts a diagnosis of pneumonia, but the patient was suffering from heart failure, and thus received the wrong treatment, which led to an injury.

1. Analysis Under Current U.S. Tort Liability Law

In this Scenario, the patient could potentially sue the manufacturer under product liability law and/or the individual or organizational healthcare providers under fault-based negligence law.

a. Manufacturer Liability

The product liability analysis under U.S. tort liability law in this Scenario is largely similar to the analysis discussed in Scenario 1.[243] First, assuming that the AI is considered a product, a claim against the manufacturer for product liability will depend on the reason for the AI's wrong prediction. If, for example, the AI predicted an incorrect pneumonia diagnosis because the AI was either improperly designed to use a data set that was too small or improperly validated using the wrong patient population, then the manufacturer might be liable for a design defect. Additionally, if the AI was designed to use a data set that was different than the data set actually used to train it, causing the final AI to operate differently than its intended design, then the manufacturer might be liable for a manufacturing defect. Finally, if the manufacturer failed to comply with monitoring, warning, or updating requirements, it might be liable for a marketing defect. On the other hand, if the AI was trained, validated, and functioning as designed to use a noninterpretable algorithmic reasoning to provide a prediction, the mere fact that the pneumonia prediction for this patient was incorrect is likely not enough to deem the product defective or the manufacturer negligent under the current product liability framework under

---

[242] *See* Sarah Jabbour et al., *Combining Chest X-Rays and Electronic Health Record (EHR) Data Using Machine Learning to Diagnose Acute Respiratory Failure,* 29 J. AM. MED. INFORMATICS ASS'N 1060, 1060 (2022) (discussing that machine learning models could enhance medical diagnosis by assisting in diagnostic evaluation of patients with acute respiratory failure).
[243] *See supra* Section IV.A.1.a.

the same risk-utility considerations and potential defenses discussed in Scenario 1. However, in contrast to Scenario 1, since a physician directly oversees this non-autonomous system, an affirmative LID defense is more likely to be successful in this Scenario than in Scenario 1.[244]

b. Healthcare Provider Liability

For a patient to successfully sue a healthcare provider for medical malpractice in this case, they must first prove that the provider breached the standard of care. With regard to both individual and organizational healthcare providers, if we assume that they complied with all the standards, guidelines, and instructions, related to the use, implementation, and oversight of the AI, for example, properly selecting the AI for use in emergency care, ensuring proper training, instruction, oversight by competent staff, properly maintaining software, and having proper policies related to the AI's use, the provider will likely not be negligent for using the AI in emergency care.

Similarly, the physician who relied on the AI's pneumonia prediction may not be negligent. Here, the AI predicted the cause of acute respiratory failure using complex algorithmic reasoning that is neither transparent nor interpretable to the human physician, making it impossible for them to understand the true basis for the AI's prediction.[245] Further, although the physician oversaw the AI and was officially tasked with making the final "medical decision," they were not able to independently assess the accuracy of the AI's prediction prior to making the decision to treat the patient for pneumonia rather than heart failure.[246] As a result, reliance on the AI's pneumonia prediction may not violate an applicable duty of care if the physician used this legally-marketed AI in accordance with its intended use and did not know and could not have known that the AI's prediction was incorrect.

---

[244] *See supra* Section II.A.4.

[245] *See* Duffourc & Gerke, *supra* note 203, at 2 (discussing how some medical black-box AI systems are noninterpretable).

[246] Of course, it is also possible that the human provider is not capable of independently assessing the accuracy of the AI's decision at any point in time. This is particularly true if the AI provides an output that exceeds current human knowledge.

2. *Analysis Under the Current and Newly Proposed EU Liability Framework*

In this Scenario, the patient could potentially sue the manufacturer for product liability and/or the individual or organizational healthcare providers for fault-based medical liability under Member States' national law, which incorporates the Current PLD, and, once enacted and transposed into national law, would incorporate the Proposed PLD and Proposed AILD.

a. Manufacturer Liability

The analysis of product liability for this non-autonomous black-box AI under the Current PLD and Proposed PLD and Proposed AILD is largely the same as the analysis in Scenario 1.[247] However, we highlight a few specific points for this Scenario. First, even under the Proposed PLD's new test for determining defectiveness, the AI's inaccurate pneumonia prediction may not be considered a defect if it was the result of the AI's black-box algorithmic reasoning process, which the manufacturer could not have been able to predict or control.[248] Second, the fact that this AI is non-autonomous may further remove it from the manufacturer's control and thus decrease the likelihood of finding a defect.[249] Third, if the manufacturer complied with all mandatory safety requirements, the likelihood of considering the AI's inaccurate pneumonia prediction to be a defect under the Proposed PLD decreases further.

b. Healthcare Provider Liability

As in Scenario 1, the patient will only be able to recover if they can prove fault of a healthcare provider.[250] As in the United States, if both the individual and organizational healthcare providers' use, implementation, and oversight of the AI did not violate any legal duties of care, it will be difficult for a patient to prove fault in connection with the provider's use of the AI in emergency care.[251]

---

[247] *See supra* Section IV.A.2.a.
[248] *See* Proposed PLD, *supra* note 14, art. 10(1)(e).
[249] *See id.* art. 6(1)(b), (e), (h) (providing considerations for determining whether a product is defective).
[250] *See supra* Section IV.A.2.b.
[251] *See supra* Section IV.A.2.b.

Additionally, and again similar to the situation in the United States, if this AI makes predictions by recognizing patterns in data that the human brain cannot discern or provides a time-sensitive prediction that the physician cannot make themselves, the physician's failure to recognize the incorrect prediction may not be considered a breach of a duty of care if they could not have known that the AI's pneumonia prediction was incorrect at the time they relied on it to prescribe the emergency medical treatment.[252]

Also, as discussed in Scenario 1, if the plaintiff is able to prove an individual or organizational provider's fault, the Proposed AILD may provide them with an evidentiary presumption that the provider's fault caused the AI's incorrect pneumonia prediction.[253] However, this is only the case if the provider's fault preceded the AI's prediction, which would occur, for example, in this Scenario if the hospital was at fault by deciding to use the AI in the wrong patient population. Because the EC determined that claimants will not encounter the burden of proof difficulties in cases when "AI systems providing medical analysis or even suggestions for diagnosis and treatment, which are feeding into a decision on diagnosis and treatment, but [when] that decision is ultimately taken by a human physician," the Proposed AILD's evidentiary rules will not assist the patient if there is no fault that *preceded* the AI's inaccurate pneumonia prediction.[254] The EC's approach is based on an assumption that it is always possible for the doctor who prescribed the pneumonia treatment to independently assess the accuracy of the AI's pneumonia prediction prior to providing emergency treatment. Of course, as discussed above, this may not be possible if the AI uses reasoning that exceeds human intelligence or when the doctor relied on the AI for time-sensitive decision making. In such situations, there may be no fault to trigger the Proposed AILD's evidentiary presumption on causation.

---

[252] Duffourc & Gerke, *supra* note 203, at 4 (discussing how there may be no breach of a duty of care by a healthcare provider who reasonably relies on the output of a legally-marketed noninterpretable black-box AI).

[253] *See* Proposed AILD, *supra* note 15, art. 4(1) (stating that "national courts shall presume . . . the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output," where certain conditions are met).

[254] Commission Staff Working Document, *supra* note 182, § 2.2.

C. Scenario 3: Direct-to-Consumer Black-Box AI

Imagine that a consumer, without consulting a doctor, turns to a DTC general-use black-box AI to answer the following health-related question at 9:00 AM:

> Question:
>
> I'm a 50-year old man that weighs 180 pounds. I take medication for high cholesterol and high blood pressure. I'm having heartburn, nausea, and shoulder pain. What might be wrong with me?
>
> Response:
>
> I'm sorry to hear that you're experiencing these symptoms. As an AI language model, I am not qualified to provide medical diagnoses or treatment advice. However, based on your symptoms, there are a few possible causes:
>
> 1. Acid reflux: The heartburn and nausea could be symptoms of acid reflux, which occurs when stomach acid flows back into the esophagus. This can cause a burning sensation in the chest, as well as nausea and vomiting.
>
> 2. Heart attack: Shoulder pain can sometimes be a symptom of a heart attack, especially if it is accompanied by chest pain, shortness of breath, and sweating. Given your age and medical history, it's important to take this possibility seriously and seek medical attention immediately if you experience these symptoms.
>
> 3. Gallbladder problems: Shoulder pain can also be a symptom of gallbladder problems, such as gallstones or inflammation of the gallbladder. This can cause nausea and vomiting as well.
>
> It's important that you see a doctor as soon as possible to determine the cause of your symptoms and receive appropriate treatment. Your

> doctor may perform tests such as an electrocardiogram (ECG), blood tests, or imaging tests to help diagnose the underlying condition.[255]

The user in this Scenario interpreted the AI's numbered list as an indication that the first diagnosis listed was the most likely diagnosis. This put the user at ease, and they took some medication for heartburn and planned to visit the doctor during their afternoon work break, which for them was "as soon as possible." At 11:00 AM, the user suffered an injury from a heart attack.

1. Analysis Under Current U.S. Tort Liability Law

In this Scenario, the consumer may be able to sue the manufacturer under strict or fault-based product liability law. Since there are no healthcare providers involved in this Scenario, there is no potential fault-based liability based on an individual or organizational healthcare provider's actions. Thus, the following analysis focuses exclusively on manufacturer liability.

Injuries caused by DTC general-use AI that provides medical information and advice also present challenges to the traditional product liability framework. Assuming that this DTC general-use black-box AI is a product subject to product liability law, the analysis of a *design defect* claim is largely similar to Scenarios 1 and 2.[256] However, unlike the health-specific black-box AI in Scenarios 1 and 2, this general-use AI is not likely to be considered an unavoidably dangerous product in a design defect claim because it is not a drug or medical device.[257] As a result, it is less likely to pass a risk-utility test, particularly if it could have been better designed to avoid misleading consumers into believing that they are receiving personalized and reliable medical information and advice, for example, by refraining from providing answers to users' questions that seek an individualized medical diagnosis.

As discussed in Scenarios 1 and 2, the consumer may be able to prove a *manufacturing defect* if the data used to train the general-use AI was not the training data specified by the manufacturer's design.[258] However, this will likely not be the case for a general-use AI that is designed to use vast amounts of data

---

[255] *Shoulder Pain and Heartburn*, OPENAI (July 17, 2023), https://perma.cc/8C6J-2N57 (generated using ChatGPT-3.5 May 24 version).
[256] *See supra* Sections IV.A.1.a, IV.B.1.a.
[257] *See supra* Section II.A.4.
[258] *See supra* Sections IV.A.1.a, IV.B.1.a.

available on the internet generally rather than a health-specific AI that was designed, trained, and validated for a health-specific purpose.

A *failure to warn* claim may offer some protection to the consumer in this case. Warning users about its lack of medical expertise may be sufficient to warn users of the risks of relying on the AI's answers to health-related questions.[259] However, despite this AI's admission that it is "not qualified to provide medical diagnoses or treatment advice," it continued to do just that.[260] A court may find that simply disclaiming that the information it provides is not medical advice does not magically exempt it from being medical advice or, at least, misleading the user into thinking that it is medical advice.[261] In fact, the recommendation to seek medical advice itself may be considered medical advice.[262] If the user's reliance on the AI's medical advice, despite the AI's warning, is a reasonably foreseeable misuse of the AI or if the manufacturer became aware of an unreasonable risk for users who relied on the AI's answers and failed to update its warnings, the manufacturer may be liable for a marketing defect.[263]

Finally, we note that even if this AI is considered defective, the manufacturer may be able to avoid strict product liability by proving that the user either (1) misused the system by relying on medical advice or information "in direct contravention of the product's warnings and instructions" if this misuse is not reasonably foreseeable, or (2) had subjective knowledge of the specific risks associated with relying on the AI's medical information or advice

---

[259] *See Will ChatGPT Transform Healthcare?*, *supra* note 65, at 505 (noting that warnings may help protect consumers).

[260] *Shoulder Pain and Heartburn*, *supra* note 255.

[261] *See* Bennett v. Bureau of Pro. & Occupational Affs., 214 A.3d 728, 737 (Pa Commw. Ct. 2019) (noting that a disclaimer does not cure misleading representations); *cf.* Khan v. BDO Seidman, LLP, 948 N.E.2d 132, 154 (Ill. App. Ct. 2011), *aff'd sub nom*, Khan v. Deutsche Bank AG, 978 N.E.2d 1020 (Ill. 2012) (finding boilerplate disclaimer of no fiduciary relationship insufficient when party served as broker and gave investment advice); *In re* Hodge, 407 P.3d 613, 644 (Kan. 2017) (finding letter stating that an attorney who was not providing legal advice did not preclude their legal advice from being legal advice).

[262] Dr.'s Co. v. McDonough, No. 01-00-00741-CV, 2002 WL 2024260, at *3 (Tex. Ct. App. Aug. 30, 2002); *see also* Ghee v. USAble Mut. Ins. Co., No. 1200485, 2023 WL 2720996, at *8–*9, *11 (Ala. Mar. 31, 2023) (finding that the defendant was not entitled to ERISA preemption because "advis[ing] [participant] to return to the hospital" could be considered negligent medical advice).

[263] *See supra* Section II.A.3.

and chose to assume that risk.[264] The manufacturer may also be able to avoid or reduce their liability by showing that the consumer's reliance on the AI's output was objectively unreasonable (or negligent) in either negligent product liability cases or in jurisdictions that allow a comparative allocation of fault in strict liability cases.

2. Analysis Under the Current and Newly Proposed EU Liability Framework

In this Scenario, the patient could potentially sue the manufacturer for product liability under Member States' national law, which incorporates the Current PLD and would incorporate the Proposed PLD and Proposed AILD (once enacted and transposed into national law). Again, since no healthcare providers are involved in this Scenario, there is no potential fault-based liability under Member States' national law based on an individual or organizational healthcare provider's actions. Thus, the following analysis focuses exclusively on strict and fault-based manufacturing liability.

Under the Current PLD, this DTC AI is less likely to be considered a product subject to strict product liability because it is probably not embedded into a tangible product.[265] Since the Proposed PLD would consider this AI a product, it provides a better basis for the plaintiff to recover under strict liability framework. The AI in this Scenario presents the user with a sophisticated, seemingly individualized response to a user-specific health question, which a court might weigh in finding a defect under the product's presentation factor.[266] Additionally, its instruction to the user "to see a doctor as soon as possible" might also weigh in favor of finding a defect if the user's interpretation of this instruction, that is to see a doctor during a work break hours later, is considered reasonable. Finally, if this is an adaptive DTC AI, its ability to continuously learn will also weigh in favor of finding it defective.

On the other hand, under the Proposed PLD, if the court finds that following the AI's instruction would have required a different response by the user—that

---

[264] *See supra* Section II.A.4.

[265] *See supra* Section III.A.1. (discussing whether software would be considered a movable under the Current PLD); *see also* Current PLD, *supra* note 148, art. 2 (discussing scope of the Directive as applying only to "movables" with limited exceptions).

[266] *See* Current PLD, *supra* note 148, art. 6(1)(a) (listing presentation of a product as a factor for determining defectiveness); Proposed PLD, *supra* note 14, art. 6(1)(a); *see also supra* Sections III.A.1 and II.B.1.

is, immediate consultation with a human doctor—this might weigh against finding a defect. Additionally, the user's reliance on the AI's response to this medical question, despite the AI's warning that it is not qualified to provide medical advice, might be considered a reasonably foreseeable misuse of the product weighing in favor of finding a defect.[267] However, if the court finds that the user's reliance on the AI's output, despite its disclosure that it cannot provide medical advice, was unreasonable, the AI may not be considered defective in this case.[268] Additionally, the product's ability to continually learn—if not locked—will likely weigh in favor of finding a defect.[269] Finally, if the manufacturer complied with all mandatory safety requirements, this would likely weigh against finding a defect.[270]

A negligent product liability claim under Member States' national fault-based liability law would likely involve similar considerations when judging whether the manufacturer's behavior in designing, manufacturing, and marketing this DTC AI breaches a duty of care.[271] For example, if the manufacturer's failure to design the system to refrain from providing any response to a user's request for individualized medical advice is considered unreasonable, a manufacturer may face liability under a Member's State's fault-based liability law. Finally, the Proposed AILD would provide the injured user with an evidentiary presumption that the manufacturer's fault caused the AI's given medical advice in response to the user's question, but the user would still have to prove that the AI's response caused, or partially caused, their injury under the Member States' burden of proof rules.[272]

---

[267] *See* Proposed PLD, *supra* note 14, art. 6(1)(b) (listing "reasonably foreseeable use and misuse of the product" as a factor for determining defectiveness); *see also supra* Section III.B.1.

[268] *See* Proposed PLD, *supra* note 14, art. 6(1)(b); *see also supra* Section III.B.1.

[269] *See* Proposed PLD, *supra* note 14, art. 6(1)(c) (listing the "effect on the product of any ability to continue to learn after deployment" as a factor for determining defectiveness); *see also supra* Section III.B.1.

[270] *See* Proposed PLD, *supra* note 14, art. 6(1)(f) (listing "product safety requirements" as a factor for determining defectiveness); *see also supra* Section III.B.1.

[271] *See supra* Section III.A.1.

[272] *See* Proposed AILD, *supra* note 15, art. 4(1) (stating that "national courts shall presume . . . the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output," where certain conditions are met); *id.* recital 15 (discussing the scope of the Directive); *see also supra* Section III.B.2.

V. LESSONS LEARNED FROM THE EUROPEAN UNION'S APPROACH TO BLACK-BOX AI LIABILITY

A comparison of the EU and U.S. liability frameworks for medical injuries caused by black-box AI systems and the application of these frameworks in the specific Scenarios involving black-box AI systems reveals several lessons that the United States can learn from the European Union's approach to AI liability, and specifically, liability for medical injuries caused by black-box AI systems. This Part first argues for the relevance of the European Union's approach and then draws lessons for the United States.

A. Relevance of the European Union's Approach

While we recognize that the jurisdictions may differ on important legal and policy positions that can influence the adoption, use, regulation, and liability of black-box AI (i.e., discovery, general AI regulations, privacy, access to health care, etc.), we nevertheless find that the United States can glean important lessons from the European Union's approach to the common challenges that black-box AI poses to the liability systems in both jurisdictions.

As demonstrated in Sections II to IV, remarkably similar principles govern liability for manufacturers and healthcare providers in the United States and European Union. While they both recognize fault-based (or negligent) manufacturer liability, both jurisdictions also take a stricter approach to liability for injuries caused by defective products. Both jurisdictions also consider consumer expectations and risk/utility factors to judge whether a product is defective in design.

With regard to healthcare provider liability, both jurisdictions take fault-based approaches and will generally only impose liability when the individual or organizational healthcare provider breached a duty of care. Although there are some differences in the way the jurisdictions (even within the European Union) approach evidentiary issues, including the burden and standard of proof, the common core of strict product liability for manufacturers and fault-based medical liability for healthcare providers remains. This, combined with common liability challenges that black-box AI poses in both jurisdictions, exemplified by the comparative analysis of the hypotheticals in Section IV, provides a compelling reason to pay attention to the European Union's approach to liability for AI-caused injuries.

The European Union is certainly ahead of the United States when it comes to initiatives to update the legal framework to address new liability issues raised

by AI. Generally, the EC recognizes that "[it] is important that victims of accidents of products and services including emerging digital technologies like AI do not enjoy a lower level of protection compared to similar other products and services, for which they would get compensation under national tort law."[273] The Proposed PLD and Proposed AILD represent the EC's attempts to remedy potential liability gaps and ensure legally trustworthy AI through ex-post liability rules.

The European Union's general approach to AI liability in the proposed Directives addresses some of the challenges that arise with the use of black-box AI for healthcare, but it also falls short in some regards. An analysis of both the successes and the shortcomings of the European Union's approach, when applied in the healthcare context, provides valuable lessons for the United States. In terms of successes, both proposed EU Directives make progress toward the European Union's goal of ensuring trustworthy AI by (1) "directly addressing unique risks posed by AI" in their respective no-fault and fault-based liability approaches and (2) "mitigating the information asymmetry" through increased access to evidence and decreased burdens of proof for claimants.[274] As for shortcomings, the Proposed PLD and Proposed AILD in combination with national law still fail to provide a workable liability framework for some medical injuries caused by black-box AI. This is because the AI may not be considered defective under the Proposed PLD, and there may be no breach of a duty of care by a legally responsible person upon which to base liability under the Proposed AILD.[275] As shown in Section IV, similar gaps might emerge in the United States under the existing tort framework governing manufacturer and healthcare provider liability.

B. Lessons Learned from the European Union

The European Union's recent attempts to preemptively tackle challenges related to liability for AI-caused injuries provide four valuable lessons for U.S. stakeholders involved with the creation and use of black-box AI as they

---

[273] Report from the Commission, *supra* note 182, § 3.
[274] Duffourc & Gerke, *supra* note 203, at 1.
[275] See *id.* for more information.

inevitably face a case involving liability for a medical injury caused by this rapidly developing technology.

Lesson 1: A broad approach to AI liability fails to provide solutions to some challenges posed by black-box AI in healthcare.

Both countries face the colossal task of adapting legal systems to accommodate the new demands that accompany the rapid development and integration of AI technology in various sectors of society. The European Union is on the right path to addressing AI-related safety concerns through both ex-ante regulation and ex-post liability rules, but the diversity of AI architecture and AI applications demands a more dynamic approach to liability for AI-caused injuries, particularly in the health sector. The European Union's Proposed PLD and Proposed AILD follow the Proposed AI Act, which seeks to protect fundamental rights and ensure safety in connection with AI through ex-ante regulation. The combination of regulatory and liability-based approaches to AI reflects the EC's view that "[s]afety and liability are two sides of the same coin: they apply at different moments and reinforce each other."[276] While this general approach to AI regulation and liability is commendable, the European Union's failure to consider the unique risks posed by black-box AI in different industries, e.g., healthcare versus autonomous vehicles, leads to broad regulatory and liability rules that often fail to address crucial risks and leave worrying liability gaps.[277]

Lesson 2: Traditional concepts of human fault pose significant challenges in cases involving black-box AI.

In both jurisdictions, it may be difficult for claimants to connect AI-caused injuries to the fault of a party that can be held legally responsible under the current liability framework. While the European Union has recognized that "[t]he processes running in AI systems cannot all be measured according to duties of care designed for human conduct," the Proposed AILD fails to address the largest obstacle to fault-based liability for black-box AI: proving fault

[276] Proposed AILD, *supra* note 15, Explanatory Memorandum.

[277] For more information regarding liability gaps in the European Union, see Duffourc & Gerke, *supra* note 203, at 1.

vis-à-vis the non-compliance with an existing duty of care.[278] Under the Proposed AILD, unless a presumption of fault arises from a defendant's failure to disclose court-ordered evidence, the claimant still bears the burden of proving that a natural or legal person failed to comply with a duty of care under Member States' national law that is "reasonably likely" to have influenced an AI's injury-causing output.[279] Unfortunately, the opacity, complexity, and/or autonomy of black-box AI can cause damage that could not have been predicted or prevented by a legal person's actions or omissions, completely severing the "fault" of the AI from potentially legally responsible persons associated with its use.

Lesson 3: Product liability frameworks must consider the unique features of black-box AI.

In both jurisdictions, black-box AI challenges the traditional product liability framework because it does not always present and function like a traditional product. First, black-box AI may not be considered a product subject to product liability claims. The European Union solved this problem by explicitly clarifying that software is a "product" subject to liability under the Proposed PLD.[280]

Second, unlike traditional products, manufacturers may be able to maintain control and influence over a black-box AI's behavior following market entry, making market entry as the temporal reference to judge a product's defectiveness unsuitable for some black-box AI. In the European Union, the Proposed PLD recognizes the ability of software, including black-box AI, to change after market entry by expanding the scope of liability to include products still under the "manufacturer's control" rather than focusing solely on market entry.[281] The Proposed PLD also integrates the continuous self-learning

---

[278] EXPERT GRP. ON LIAB. & NEW TECHS., EUR. COMM'N, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES 23 (2019), https://perma.cc/PFP5-TSPN; *see* Proposed AILD, *supra* note 15, recital 24 (relying on fault as established by Member State law).

[279] *See* Proposed AILD, *supra* note 15, art. 4(1)(b) (providing a rebuttable presumption for causation when it is "reasonably likely, based on the circumstances of the case, that the fault has influenced the output produced by the AI system or the failure of the AI system to produce an output").

[280] *See* Proposed PLD, *supra* note 14, art. 4(1) (defining "product").

[281] *See id.* art. 4(5) (defining "manufacturer's control").

characteristic of adaptive algorithms as a characteristic affecting whether the product is considered "defective."[282]

Finally, it can be difficult to determine when a black-box AI is defective under the traditional framework. The Proposed PLD's test for defectiveness makes some progress toward addressing the unique capabilities of AI, but it still fails to provide workable criteria for determining product "defectiveness" for some black-box AI technologies. The Proposed PLD hinges defectiveness on a finding that the product "does not provide the safety which the public at large is entitled to expect, taking all circumstances into account."[283] The Proposed PLD also lists circumstances, including the ability of algorithms to continuously learn, which might weigh in favor of finding black-box AI defective.[284] On the other hand, it considers the manufacturer's control over the product, compliance with safety requirements, and expectations of end-users, all of which may weigh against finding a black-box AI defective.[285] Even when a product is considered defective, the Proposed PLD provides an exemption to liability when the manufacturer could not have discovered the defect under the "objective state of scientific and technical knowledge."[286]

Lesson 4: Evidentiary rules should address the difficulties that claimants will face in obtaining evidence and proving causation.

In both jurisdictions, it may be difficult for plaintiffs to obtain evidence and satisfy evidentiary burdens in some cases involving an AI-caused injury because the technical inner workings of the black-box AI are highly complex and opaque. The Proposed PLD's evidentiary rules can assist claimants with both obtaining evidence from manufacturers and proving defectiveness in cases involving complex black-box AI.[287] The Proposed AILD's evidentiary rules also address

---

[282] *See id.* art. 6(1)(c) (listing "the effect on the product of any ability to continue to learn after deployment" as a factor for determining defectiveness).
[283] *Id.* art. 6(1).
[284] *See id.* art. 6(1)(c) (listing "the effect on the product of any ability to continue to learn after deployment" as a factor for determining defectiveness).
[285] *See id.* art. 6(1)(e), (f), (h) (listing "where the manufacturer retains control over the product . . . , the moment in time when the product left the control of the manufacturer," "product safety requirements," and "specific expectations of the end-users for whom the product is intended" as factors for determining defectiveness).
[286] *Id.* art. 10(1)(e).
[287] *See id.* arts. 8(1), 9(4), 9(2) (describing presumptions of defectiveness and evidentiary rules).

some of the concerns about information asymmetry by making evidence more accessible in claims involving high-risk systems and the burden of proof less demanding in many claims involving any type of AI.[288]

CONCLUSION

This Article is the first to provide a comprehensive roadmap of liability for black-box AI in healthcare in the United States and European Union and use comparative legal methodology to identify important lessons for the United States from the newly proposed EU approach. In Part I, we explained crucial features of black-box AI and outlined the landscape of current and future uses of black-box AI in the health sector. Part I identified how and why black-box AI technology presents both benefits and risks when used for health-related purposes and described that these benefits and risks are unique because they stem from unique features of the AI's design that are not typical of traditional consumer products and medical devices. As a result, medical injuries caused by black-box AI challenge the traditional liability framework in both the United States and European Union.

In Part II, we identified the current framework that would govern liability for manufacturers and individual and organizational healthcare providers in cases involving medical injuries caused by black-box AI in the United States. When considering the current framework in the context of the unique risks presented by black-box AI systems, we revealed that there are critical liability concerns that manifest when some black-box AI systems cause medical injuries. We identified three valuable takeaways: (1) black-box AI challenges the legal standards by which products are deemed defective, (2) injuries caused by black-box AI cannot always be sufficiently connected to the fault of a legally responsible party, and (3) claimants who suffer a medical injury involving black-box AI may face significant evidentiary struggles.

In Part III, we identified the current and proposed framework governing liability for manufacturers and individual and organizational healthcare providers in cases involving medical injuries caused by black-box AI in the European Union. Laying out the current framework in the context of black-box AI revealed that the European Union and United States encounter

---

[288] Proposed AILD, *supra* note 15, arts. 3, 4.

similar liability challenges when faced with the unique risks posed by black-box AI in large part because both jurisdictions' liability frameworks are grounded in similar principles. We set forth the European Union's new approach to liability for AI-caused injuries in the Proposed PLD and Proposed AILD and analyzed these proposed liability rules in the context of black-box AI. We concluded the European Union's approach to liability addresses (and fails to address) some of the liability challenges that both jurisdictions face as a result of black-box AI in healthcare, which entails important lessons for the United States.

In Part IV, we discovered valuable insights from the European Union's approach to AI liability by analyzing liability in three hypothetical scenarios involving a medical injury caused by black-box AI under both current U.S. and current and newly proposed EU law. The comparative analysis revealed that the European Union's proposed Directives successfully tackle some common liability challenges by (1) including AI in the scope of its definition of a product in the Proposed PLD, (2) considering the dynamic nature of AI-driven products when determining whether a product is defective under the Proposed PLD, and (3) adjusting evidentiary rules in some cases to help claimants who suffer AI-caused injuries under the Proposed PLD and Proposed AILD. However, our analysis also revealed that liability gaps can manifest in both jurisdictions when a medical injury results from a healthcare provider's use of an autonomous black-box AI (Scenario 1), a non-autonomous black-box AI (Scenario 2), and a consumer's use of a general black-box generative AI (Scenario 3). We concluded that these gaps are likely to occur when the AI's noninterpretable reasoning process produces an injury-causing output, and the following are true: (1) the AI was functioning as designed by the manufacturer, (2) the manufacturer complied with safety regulations relating to the AI's development and marketing, and (3) individual and organizational healthcare providers were reasonable in their use, selection, and implementation of the AI and could not have reasonably known that the AI's output was incorrect. We also concluded that black-box AI's complex architecture and operation present evidentiary challenges for injured claimants even with the newly proposed EU approach, particularly in meeting their burden of proving fault and causation.

In Part V, we identified lessons learned from the European Union's approach to AI liability in the context of medical injuries caused by black-box AI. We found that the European Union's approach is relevant to the United States because the liability frameworks in both jurisdictions operate under similar legal principles and face common legal challenges when confronted with

black-box AI in the healthcare domain. We argued that four main lessons can be learned from the European Union's approach that are relevant for stakeholders in the United States, who will likely encounter these common challenges. First, a broad approach to AI liability fails to provide solutions to some challenges posed by black-box AI in healthcare. Second, traditional concepts of human fault pose significant challenges in cases involving black-box AI. Third, product liability frameworks must consider the unique features of black-box AI. Fourth, evidentiary rules should address the difficulties that claimants will face in cases involving medical injuries caused by black-box AI.

The possible benefits and consequences of black-box AI in healthcare are significant. Now is the time for stakeholders in the United States to give serious consideration to the potential liability associated with the use of black-box AI in the health sector. Understanding the current liability framework and its limitations is crucial to ensure consumer and patient safety and to encourage the continued development and deployment of beneficial black-box AI in healthcare. This understanding should also consider the European Union's newly proposed approach to liability for medical injuries for black-box AI to glean important lessons from the European Union's attempts to solve similar liability challenges.

ACKNOWLEDGEMENTS

This work was funded by the European Union (Grant Agreement no. 101057321). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Sara Gerke also reports grants from the European Union (Grant Agreement no. 101057099), the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the National Institutes of Health Office of the Director (NIH OD) (Grant Agreement no. 3R01EB027650-03S1 and no. 1R21EB035474-01), and the National Institute on Drug Abuse (NIDA) and the National Institutes of Health (NIH) (Grant Agreement no. 1U54DA058271-01).