

The oversight of autonomous artificial intelligence: lessons from nurse practitioners as physician extenders

Walker Morrell^{1,†,*}, Carmel Shachar^{2,‡} and
Anthony P. Weiss^{3,4,**}

¹Harvard Medical School Center for Bioethics, Boston, MA, USA

²Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School, Cambridge, MA, USA

³Beth Israel Deaconess Medical Center, Boston, MA, USA

⁴Harvard Medical School, Boston, MA, USA

*Corresponding author. E-mail: wmorrell@jd25.law.harvard.edu

ABSTRACT

The development of autonomous artificial intelligence (A-AI) products in health care raises novel regulatory challenges, including how to ensure their safety and efficacy in real-world settings. Supplementing a device-centered regulatory scheme with a regulatory scheme that considers A-AI products as a ‘physician extender’ may improve the real-world monitoring of these technologies and produce other benefits, such as increased access to the services offered by these products. In this article, we review the three approaches to the oversight of nurse practitioners, one type of physician extender, in the USA and extrapolate these approaches to produce a framework for the oversight of A-AI products. Under the framework, the US Food and Drug Administration would evaluate A-AI products and determine whether they are allowed to operate independently of physician oversight; required to operate under some physician oversight via a ‘collaborative protocol’ model;

† Walker Morrell is a graduate of the Master of Bioethics program at Harvard Medical School and an incoming student at Harvard Law School.

‡ Carmel Shachar is the executive director of the Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School.

** Anthony P. Weiss is the chief medical officer at Beth Israel Deaconess Medical Center and an associate professor of Psychiatry at Harvard Medical School.

or required to operate under direct physician oversight via a ‘supervisory protocol’ model.

KEYWORDS: autonomous, artificial intelligence, FDA, medical device, nurse practitioner, oversight

I. INTRODUCTION

Artificial intelligence (AI) is a powerful tool in health care. Although presently used primarily to provide information and decision support to health care professionals and patients, AI technologies are increasingly able to autonomously (ie without input or support from a human clinician) diagnose and treat patients.¹

Regulators are developing frameworks to ensure the safety and efficacy of AI products in health care. Currently, most regulators take a device-centered approach, with the understandable logic that AI products most closely resemble medical devices and should be regulated as such. For example, in April 2019, the US Food and Drug Administration (FDA) released a discussion paper titled, ‘Proposed Regulatory Framework for Modification to Artificial Intelligence/Machine Learning-Based Software as a Medical Device’.² In January 2021, FDA released the ‘Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device Action Plan’.³ Both the discussion paper and the action plan draw heavily from existing device-centered regulatory pathways. They highlight that any regulatory framework for the oversight of AI technologies must include mechanisms for real-world performance monitoring (ie the collection of data that indicates whether a product produces safe, effective, and reliable outcomes in real-world settings).⁴ Data on patient health status and the delivery of health care may be collected from electronic health records, claims and billing activities, and other sources.⁵ FDA is currently working to determine what reference data could be used to monitor the performance of AI technologies and how end-user feedback can be incorporated into the evaluation of these technologies.⁶

There are several challenges that arise from using this device-centered regulatory approach to govern autonomous artificial intelligence (A-AI) products, particularly associated with real-world performance monitoring. First, it could be difficult to identify data that reflects the efficacy of an A-AI product in the real world. In contrast to

1 Kun-Hsing Yu et al., *Artificial Intelligence in Healthcare*, 2 NAT. BIOMED. ENG. 719 (2018); Fei Jiang et al., *Artificial Intelligence in Healthcare: Past, Present and Future*, 2 STROKE VASC. NEUROL. 230 (2017); U.S. Food and Drug Admin., *Clinical Decision Support Software: Draft Guidance for Industry and Food and Drug Administration Staff* (2019), <https://www.fda.gov/media/109618/download> (last accessed July 11, 2022); Amber A. van der Heijden et al., *Validation of Automated Screening for Referable Diabetic Retinopathy with the IDx-DR Device in the Hoorn Diabetes Care System*, 96 ACTA OPHTHALMOL. 63 (2018).

2 U.S. Food and 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* (2019), <https://www.fda.gov/media/122535/download> (last accessed July 11, 2022).

3 U.S. Food and Drug Admin., *Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan* (2021), <https://www.fda.gov/media/145022/download> (last accessed July 11, 2022).

4 *Id.* at 6; U.S. Food and Drug Admin., *supra* note 2, at 14.

5 U.S. Food and Drug Admin., *Real-World Evidence* (2020), <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence> (last accessed May 20, 2022).

6 U.S. Food and Drug Admin., *supra* note 3, at 6.

research populations, which are narrowly defined to permit inferences on the efficacy of an intervention, clinical populations are heterogenous and may not allow these inferences. Second, real-world performance monitoring may not allow regulators to sufficiently scrutinize the safety of A-AI products. FDA estimates that only 0.5 per cent of incidents related to medical devices are reported,⁷ and the processing of an adverse event report can take several months.⁸ Third, end-user feedback will have limited value given that patients are the end-users of A-AI products and cannot evaluate and report on product performance as completely as physicians. Fourth, real-world performance monitoring relies on the collection of data post-intervention and would fail to prevent ineffective or unsafe care from occurring in the first place.

One potential solution is to supplement a device-centered regulatory scheme with a regulatory scheme that considers A-AI products as a ‘physician extender’ and expands existing physician oversight of care delivery to this new technological partner. A physician could holistically consider patient demographics and comorbidities to determine that a product is effective; monitor safety more consistently than what is allowed by real-world performance monitoring; provide more valuable end-user feedback than patients; and prevent the delivery of ineffective or unsafe care. Moreover, the physician extender paradigm is suitable for A-AI products given the role that these products will likely play in health care settings. Like other physician extenders (eg nurse practitioners, physician assistants, and nurse midwives), A-AI products may be capable of exercising independent judgement in the care of patients but will likely augment, support, and extend (and not replace) the work of physicians.⁹

In this paper, we review the three approaches to the oversight of nurse practitioners in the USA (Section II) and extrapolate these approaches to create a framework for the oversight of A-AI products (Section III). Section IV explains FDA’s proposed role. Section V discusses the potential benefits of the framework, and Section VI identifies areas for future work to implement this regulatory approach.

II. NURSE PRACTITIONER OVERSIGHT

Nurse practitioners are regulated at the state level. Although there is some level of state-to-state variation, each state’s approach generally falls into one of three categories: (i) nurse practitioners are allowed to practice independently (ie without physician involvement or oversight), (ii) nurse practitioners are required to practice in collaboration with a physician, and (iii) nurse practitioners are required to practice under the direction or supervision of a physician.¹⁰ The American Association of Nurse Practitioners (AANP) uses the terms full practice, reduced practice, and restricted practice to refer to these three approaches, respectively.¹¹

7 Amanda Craig et al., *The Need for Greater Reporting of Medical Device Incidents*, 3 EMJ INNOV. 56 (2019).

8 Daniel B. Kramer et al., *Security and Privacy Qualities of Medical Devices: An Analysis of FDA Postmarket Surveillance*, 7 PLoS ONE e40200 (2012).

9 Yu et al., *supra* note 1, at 722, 728.

10 Ann Ritter & Tine Hansen-Turton, *The Primary Care Paradigm Shift: An Overview of the State-Level Legal Framework Governing Nurse Practitioner Practice*, 20 THE HEALTH LAWYER 21, 24 (2008).

11 Am. Ass’n Nurse Pracs., *State Practice Environment*, <https://www.aanp.org/advocacy/state/state-practice-environment> (accessed Feb. 22, 2022).

Nurse practitioners in full practice states are permitted to diagnose, treat, and prescribe without physician oversight. Nevertheless, they are subject to state legislation and oversight from the state board of nursing. Regardless of whether a state takes a full, reduced, or restricted practice approach, states limit some aspects of nurse practitioner practice. These limitations are generally assigned to higher-risk activities within a nurse practitioner's scope-of-practice and activities outside a nurse practitioner's scope-of-practice.

Nurse practitioners in reduced and restricted practice states practice under the oversight of one or more physicians with the expectations and responsibilities for both parties embedded in a written agreement.¹² This co-created agreement details the activities within a nurse practitioner's scope-of-practice, the drugs that a nurse practitioner is authorized to prescribe, and the degree of physician oversight. States often identify certain topics that must be addressed in any written agreement, such as the mechanism for physician review of the nurse practitioner's prescribing practice,¹³ and/or certain items that cannot be included in any written agreement. For example, the state of Arkansas mandates that written agreements cannot include the prescription of controlled substances.¹⁴ Generally, however, states give the nurse practitioner and the physician(s) significant latitude to determine the contents of their written agreement. These written agreements are oftentimes referred to as 'collaborative agreements' in reduced practice states and supervisory 'protocols' in restricted practice states.¹⁵ This flexible approach has the benefit of allowing physicians and nurse practitioners the opportunity to tailor the scope-of-practice to reflect the realities of their capabilities, working relationship, and practice needs.

Although nurse practitioners in reduced and restricted practice states practice under the oversight of a physician, the nurse practitioner-physician relationship is fundamentally different in these two scenarios.¹⁶ Nurse practitioners and physicians in reduced practice states 'work together as professionals on equal or near-equal terms', whereas nurse practitioners in restricted practice states are considered subordinate to physicians and receive the authority to practice and prescribe through delegation by a physician.¹⁷ As a result of this difference, nurse practitioners in restricted practice states may receive more physician oversight than nurse practitioners in reduced practice states, although this can vary by state and by written agreement.

The three approaches to nurse practitioner regulation are the result of a long-standing debate regarding the extent of physician oversight necessary to ensure the safety and quality of health care delivered by nurse practitioners. Some argue that nurse practitioners are capable of delivering safe and effective care with little to no physician oversight, and furthermore, that allowing nurse practitioners to practice independently

12 Ritter & Hansen-Turton, *supra* note 10, at 24.

13 244 Mass. Code Regs. 4.07 (2021).

14 Ark. State Bd. of Nursing, *Chapter 3: Registered Nurse Practitioner*, https://www.healthy.arkansas.gov/images/uploads/pdf/Rules.Chapter03-Effective_5-15-22.pdf (2022).

15 Ritter & Hansen-Turton, *supra* note 10, at 24.

16 *Id.* at 24.

17 *Id.* at 24.

increases access to health care.¹⁸ Others argue that nurse practitioners cannot safely practice medicine without some level of physician oversight.¹⁹

There is a robust body of evidence that supports the arguments in favor of increased nurse practitioner independence. In a systematic review of the literature on nurse practitioner regulations and health care outcomes, Yang et al.²⁰ conclude that full practice regulations increase health care access and utilization, particularly for rural and underserved populations, without sacrificing the quality of care. Furthermore, some studies suggest that the cost of delivering care is lower in states with full nurse practitioner independence.²¹

III. FRAMEWORK FOR THE OVERSIGHT OF A-AI PRODUCTS

Our proposed framework is built upon three assumptions. First, we predict that allowing A-AI products to operate independently will lead to similar benefits as allowing nurse practitioners to practice independently (ie increased health care access and utilization and decreased costs). Second, we assume that some real-world physician oversight of A-AI products will be necessary to ensure their safety and efficacy, particularly when this type of product first becomes available. Third, we predict that most, if not all, A-AI products will be deployed in settings in which a physician is available to provide real-world oversight. This includes A-AI products used in traditional health care settings, such as hospitals, as well as untraditional health care settings, such as diagnostic websites or digital kiosks at chain pharmacies for which one or more physician(s) provide back end oversight.

To balance the potential benefits of independent operation and the necessity of real-world oversight, we propose a framework that includes three models for the oversight of A-AI products. These models are based upon the full, reduced, and restricted practice approaches to nurse practitioner oversight. Under this proposed framework, some A-AI products would be allowed to operate independently of physician oversight; some required to operate under some physician oversight via a ‘collaborative protocol’ model; and some required to operate under direct physician oversight via a ‘supervisory protocol’ model. FDA would be responsible for determining the necessary level of oversight when it reviews an A-AI product.

To create the framework, we reviewed the nurse practitioner regulations in Massachusetts (full practice), Arkansas (reduced practice), and Florida (restricted practice)²² and extrapolated these regulations to A-AI products. We focused our review (Table 1, left column) on the scope-of-practice and prescriptive practice regulations since these would be most similar to the operation of A-AI products. We

18 Bo Kyum Yang et al., *State Nurse Practitioner Practice Regulations and U.S. Health Care Delivery Outcomes: A Systematic Review*, 78 MED. CARE RES. REV. 183 (2021); Jeffrey Traczynski & Victoria Udalova, *Nurse Practitioner Independence, Health Care Utilization, and Health Outcomes*, 58 J. HEALTH ECON. 90 (2018).

19 Ritter & Hansen-Turton, *supra* note 10, at 22; Kevin B. O’Reilly, *3 Big Reasons Why Letting NPs Practice Independently Is a Bad Idea*, AM. MED. ASS’N, <https://www.ama-assn.org/practice-management/scope-practice/3-big-reasons-why-letting-nps-practice-independently-bad-idea> (accessed Sep. 14, 2020); *AMA Successfully Fights Scope of Practice Expansions that Threaten Patient Safety*, AM. MED. ASS’N, <https://www.ama-assn.org/practice-management/scope-practice/ama-successfully-fights-scope-practice-expansions-threaten> (last accessed July 11, 2022).

20 Yang et al., *supra* note 18, at 192.

21 Yang et al., *supra* note 18, at 191, 192.

22 Am. Ass’n Nurse Pracs, *supra* note 11.

Table 1. Framework for the real-world oversight of A-AI products

Nurse practitioner (NP) regulation	Regulatory model	A-AI product regulation
(NP 1) NPs may practice in the clinical categories for which they have ‘attained and maintained certification’. ²³	Independent	(AI 1) A-AI products are only authorized to perform the functions for which they have been permitted marketing by FDA in the populations for which they have been permitted marketing by FDA.
(NP 2) NPs are authorized to perform acts approved by the State Board of Nursing. ²⁴	Collaborative	(AI 2) A-AI products are only authorized to perform the functions for which they have been permitted marketing by FDA in the populations for which they have been permitted marketing by FDA.
(NP 3) NPs must practice in accordance with a collaborative agreement with a physician. ²⁵ The collaborative agreement must be stored on site and produced upon request by the State Board of Nursing. ²⁶		<p>(AI 3) The A-AI product must operate in accordance with a collaborative protocol with one or more physicians who work at the practice deploying the product; specialize in the same type of medicine as the product; and have received training to oversee the product.</p> <p>(AI 3a) The collaborative protocol must specify the processes for oversight of the A-AI product.</p> <p>(AI 3b) The collaborative protocol must specify the functions of the A-AI product that are toggled on/off.</p> <p>(AI 3c) The physician(s) must store the collaborative protocol on site and produce it upon request by a regulatory authority.</p>

(Continued)

excluded regulations that were not directly relevant to the real-world oversight of nurse practitioners, such as training and licensure requirements. In addition, we focused our review on the regulations in each state that set forth the basic full practice, reduced practice, and restricted practice approaches. We excluded regulatory text in Massachusetts and Florida that did not fit neatly into these states’ respective full practice and restricted practice categorizations.

23 244 Mass. Code Regs. 4.06 (2021).

24 Ark. State Bd. of Nursing, Chapter 3: Registered Nurse Practitioner, https://www.healthy.arkansas.gov/images/uploads/pdf/Rules.Chapter03-Effective_5-15-22.pdf (2022).

25 *Id.*

26 *Id.*

Table 1. Continued

Nurse practitioner (NP) regulation	Regulatory model	A-AI product regulation
<p>(NP 4) ‘If a collaborative practice results in complaints’, the state medical board may review the role of the physician and/or the NP to determine if they are ‘unable to manage [their] responsibilities under the agreement without an adverse effect on the quality of care of the patient’.²⁷</p>		<p>(AI 4) If a collaborative practice agreement results in issues or complaints, or if there is evidence that an A-AI product is operating outside its collaborative protocol, a regulatory authority may review and revoke the collaborative protocol, and, if necessary, revoke the physician’s authority to enter into a collaborative protocol with an A-AI product.</p>
<p>(NP 5) NPs may practice in the clinical categories for which they have attained and maintained certification.</p>	Supervisory	<p>(AI 5) A-AI products are only authorized to perform the functions for which they have been permitted marketing by FDA in the populations for which they have been permitted marketing by FDA.</p>
<p>(NP 6) NPs ‘must enter into a supervisory protocol with at least one physician’, and NPs may only perform those functions in the supervisory protocol.²⁸ The supervisory protocol must be stored on site and produced upon request by the State Board of Nursing.²⁹</p>		<p>(AI 6) The A-AI product must operate in accordance with a supervisory protocol with one or more physicians who work at the practice deploying the product; specialize in the same type of medicine as the product; and have received training to oversee the product.</p> <p>(AI 6a) The supervisory protocol must specify the processes for oversight of the A-AI product.</p> <p>(AI 6b) The supervisory protocol must specify the functions of the A-AI product that are toggled on/off.</p>

(Continued)

III.A. Independent Operation

A-AI products at this level would be allowed to perform the functions for which they have been permitted marketing by FDA without oversight from a physician (Table 1, AI 1). They would only be allowed to perform these functions in permitted populations.

Approval at this level would be reserved for the lowest risk products. Under this framework, there would perhaps be greater and easier access to these products because they would not need the close involvement of a physician.

27 17-87-310 Ark. Code R. § (d) (LexisNexis 2021).

28 Fla. Stat. § 464.012 (2021).

29 *Id.*

Table 1. Continued

Nurse practitioner (NP) regulation	Regulatory model	A-AI product regulation
<p>(NP 7) The supervisory physician shall ‘maintain supervision for directing the specific course of medical treatment’.³⁰</p> <p>(NP 8) Unprofessional conduct, which includes ‘practicing beyond the scope of the licensee’s license’,³¹ constitutes ‘grounds for denial of a license or disciplinary action’.³²</p> <p>N/A</p>	<p>Learning Algorithms</p>	<p>(AI 6c) The physician(s) must store the supervisory protocol on site and produce it upon request by a regulatory authority.</p> <p>(AI 7) Any diagnosis, treatment, or prescriptive decision made by the A-AI product must be reviewed by at least one supervisory physician prior to its implementation.</p> <p>(AI 8) If a supervisory protocol results in issues or complaints, or if there is evidence that an A-AI product is operating outside the supervisory protocol, a regulatory authority may review and revoke the supervisory protocol, and, if necessary, revoke the physician’s authority to enter into a supervisory protocol with an A-AI product.</p> <p>(AI 9) A-AI products that are learning algorithms will be re-reviewed by FDA and may be re-categorized to a new oversight level when:</p> <ol style="list-style-type: none"> 1. The learning algorithm experiences a software modification outside the Software as a Medical Device Pre-Specifications and the Algorithm Change Protocol; and 2. The software modification leads to a new intended use of the algorithm.

The left column presents abridged regulations for nurse practitioner oversight. The middle column indicates the oversight model. The right column presents the framework for the oversight of A-AI products.

III.B. Collaborative Oversight

A-AI products at this level would receive oversight from one or more physicians. The physician(s) would determine the amount of oversight and the processes for providing oversight and record these details in a collaborative protocol (Table 1, AI 3a). In addition, the physician(s) would customize the scope-of-practice of the A-AI product and record this in the collaborative protocol (Table 1, AI 3b). The physician(s) may allow an A-AI product to execute every diagnostic, treatment and prescriptive

30 *Id.*

31 Fla. Admin. Code Ann. R. 64B9-8.005 (2021).

32 Fla. Stat. § 464.018 (2021).

function for which it has been permitted marketing by FDA, or they may limit one or more functions of the product.

It would be the responsibility of the physician(s) to design the collaborative protocol in a way that allows them to effectively monitor the safety and efficacy of the product. Physicians may consider their capacity to monitor the product, the number of patients at a site who may be treated by the product, the vulnerability and other aspects of the patient population who may be treated by the product, and their own comfort level with allowing the AI to operate autonomously.

The physician(s) and the product manufacturer would be jointly responsible for ensuring that the A-AI product is implemented in accordance with the protocol. Manufacturers would be responsible for building features into their products to allow physicians to easily monitor the actions of the A-AI product to ensure that it complies with the collaborative protocols. FDA could ensure that manufacturers build these features into A-AI products by only permitting marketing of products that are compatible with collaborative protocols, including customizable features to reflect the content of those protocols.

Two mechanisms in the framework ensure that the physician(s) is/are appropriately positioned to meet a minimum level of oversight. First, the framework specifies that the physician(s) must work at the practice (ie hospital, nursing home, doctor's office, etc.) utilizing the A-AI product (Table 1, AI 3). Second, the framework requires that a physician can only utilize an A-AI product in a collaborative protocol if the product is within their specialty and they have received training to oversee the product (Table 1, AI 3).

In addition, there should be a regulatory authority to review and revoke a collaborative protocol and/or a physician's capacity to provide oversight via a collaborative protocol (Table 1, AI 4). Since this review would primarily focus on the physician(s) providing oversight of an A-AI product, and not on the product itself, state medical boards may be in the best position to assume this role. Provided that FDA permits marketing of safe and effective A-AI products and physicians follow the oversight requirements of the proposed framework, we would not expect this role to be overly burdensome for state medical boards, although it may require somewhat of an expansion of capacity and expertise. It remains to be seen, however, whether state medical boards would want to assume and/or be capable of executing this responsibility. An alternative or supplemental approach could be regulatory oversight by the Centers for Medicare & Medicaid Services (CMS). CMS could require that A-AI products operate in accordance with the proposed framework as a condition for billing for the use of the products. If an A-AI product is being used without a collaborative or supervisory protocol or in a way that does not align with its protocol and CMS is billed for its use, the responsible physician(s) could be held liable for defrauding the Federal Government under the civil False Claims Act.³³

Patients would benefit from increased access to the products placed in this regulatory category over the current approach. This is because physicians could better determine when they would have to be deeply involved in the use of these products, or when these products could operate with less physician oversight. For example, a

33 False Claims Act, 31 U.S.C. §§ 3729–3733.

physician could allow a product to autonomously treat patients with basic hypertension but be more involved when the product treats patients with hypertension who recently experienced a heart attack.

III.C. Supervisory Oversight

The supervisory model is similar to the collaborative model but with greater oversight and controls. The supervisory model increases oversight by requiring that the physician(s) review every diagnosis, treatment, and prescription decision made by the A-AI product prior to its implementation (Table 1, AI 7). The supervisory physician(s) may define what the review process looks like, but it is expected that the review process will ensure that every decision made by the technology meets clinical standards for safety and efficacy. If more than one physician supervises an A-AI product at this level, only one physician would be required to review each decision. In some ways, this approach is closest to the existing A-AI/physician relationship because it does not allow the A-AI product to operate autonomously.

To illustrate the difference between an A-AI product operating under a collaborative protocol as opposed to a supervisory protocol, we consider the example of a product that autonomously diagnoses ear infections and prescribes antibiotic treatment. Under a collaborative protocol, the overseeing physician may review every tenth prescription and/or review every prescription for pediatric patients to ensure the A-AI product is operating safely and effectively. Under a supervisory protocol, the overseeing physician would be required to review every diagnosis, treatment, and prescription decision made by the A-AI product.

III.D. Learning Algorithms

A-AI products that are learning algorithms, that is, algorithms that continuously adapt and change as they process data in the real-world, may evolve following FDA approval, and these changes may require a re-review and potentially a re-categorization of the level of real-world oversight. In its discussion paper on AI technologies, FDA suggests that it will require the manufacturers of learning algorithms to submit Software as a Medical Device Pre-Specifications (ie those modifications anticipated by the manufacturer) and an Algorithm Change Protocol (ie the manufacturer's methods for ensuring that pre-specified modifications are implemented safely and effectively).³⁴ Whenever a software modification occurs outside the approved pre-specifications and protocol and leads to a new intended use of the algorithm, FDA plans to execute a new pre-market review of the learning algorithm.³⁵ FDA could re-review and recategorize an A-AI product's oversight level during this pre-market review (Table 1, AI 9).

IV. FDA'S ROLE

FDA would be responsible for determining whether an A-AI product is allowed to operate independently or required to operate under collaborative or supervisory oversight. In order to implement the proposed framework, FDA needs to develop guidance on

34 U.S. Food and Drug Admin., *supra* note 2, at 10.

35 U.S. Food and Drug Admin., *supra* note 2, at 12, 13.

how to make this determination. We identify three points to consider here but do not offer a system for weighing these considerations to determine oversight level.

First, FDA should titrate the level of oversight to the risk posed by an A-AI product. In FDA's AI discussion paper, the agency references the International Medical Device Regulators Forum's (IMDRF) framework for the risk categorization of software as a medical device.³⁶ The IMDRF framework identifies two dimensions that are relevant to risk categorization—the significance of information provided by the product to the health care decision and the state of the health care situation or condition being treated.³⁷ A-AI products would be placed in the 'treat or diagnose' category of the former variable since they provide information that immediately informs diagnostic or treatment actions. Depending on whether the information is used for non-serious, serious, or critical conditions, an A-AI product may be considered Category II (medium impact), III (high impact), or IV (very high impact).³⁸

The IMDRF risk categorization framework could be used to assign minimum levels of oversight for A-AI products. Category II products could be allowed to operate independently unless other considerations lead FDA to require collaborative or supervisory oversight; Category III products could be required to operate in at least a collaborative protocol; and Category IV products could always be required to operate in a supervisory protocol.

Second, the amount of physician oversight should decrease as the quality of evidence supporting the performance of the A-AI product increases. External validation methods that mimic clinical conditions will provide the highest-quality evidence for the performance of an A-AI product. For example, consenting patients could be independently evaluated by one or more physicians and an A-AI technology. Researchers could determine the extent to which the product reaches the same conclusions as the physician(s) as well as compare the accuracy of the technology and physician recommendations.

Third, FDA should consider the representativeness of the datasets used to train and validate A-AI products. If these datasets are unrepresentative of the population for which a product is intended, the technology may produce biased outputs.³⁹ Datasets may be unrepresentative if they lack sufficient representation from patients of a certain race, ethnicity, sex, age, or other demographic factor. Datasets may also be unrepresentative if they use a proxy variable, but the proxy variable is a poor estimate of the variable of interest.

FDA is attuned to the issues of algorithmic bias and generalizability,⁴⁰ and we expect that any A-AI product permitted marketing by the agency will be thoroughly reviewed to ensure the technology has been trained and validated on representative data. With more transparency around underlying datasets, physicians could also determine if a

36 IMDRF Software as a Medical Device (SaMD) Working Group, 'Software as a Medical Device': Possible Framework for Risk Categorization and Corresponding Considerations (2014), <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf> (accessed July 11, 2022).

37 *Id.* at 12–14.

38 *Id.* at 13–15.

39 Sandeep Reddy et al., *A Governance Model for the Application of AI in Health Care*, 27 J. AM. MED. INFORM. ASSOC. 491 (2020).

40 U.S. Food and Drug Admin., *supra* note 3, at 5, 6.

particular product needs closer oversight because it does not match well that physician's particular patient population.

V. POTENTIAL BENEFITS

V.A. Flexibility

The strongest argument for moving away from a purely medical device model of regulating A-AI products to a physician extender model is that we will better be able to titrate oversight based on the particulars of each product. Some A-AI products are fairly low risk, and a framework for allowing them to operate autonomously would improve capacity within our health care system. Some products should still have an intersection with trained professionals but could do the work of physician extenders.

V.B. Increased Health Care Access and Utilization and Decreased Cost

As discussed above, there is ample literature to suggest that allowing nurse practitioners to practice independently has health care access and cost benefits. Traczynski et al.⁴¹ hypothesize that physicians in states with fully independent nurse practitioners spend less time on the supervision of nurse practitioners, allowing them and nurse practitioners to spend more time on the delivery of care, and in turn, increasing the availability of appointments. The mechanisms by which increased nurse practitioner independence leads to decreases in the cost of providing care are unclear, but the correlation may be partially explained by the fact that the reimbursement rate for nurse practitioners is 85 per cent of the reimbursement rate for physicians.⁴²

We believe the potential benefits of allowing A-AI products to operate independently or with minimum necessary oversight will be similar to the benefits of allowing nurse practitioners to practice independently, including increased access to, increased utilization of, and decreased cost of health care. More appointments may be available for the services offered by A-AI products as they increase bandwidth at practices. In addition, there may be greater access to the services not offered by these products, given that clinicians will have more time to spend on these other services. The cost of delivering health care may decrease if the reimbursement rate for services delivered by A-AI products is lower than the reimbursement rate for services delivered by human clinicians. Furthermore, there is reason to believe that the magnitude of the potential benefits of allowing A-AI products to operate independently will be greater than the magnitude of the benefits of full nurse practitioner practice. One nurse practitioner can only treat one patient at a time, whereas A-AI software could be used to simultaneously treat multiple patients in multiple locations.

As A-AI products proliferate, empirical evidence should be collected to determine whether and how these benefits are realized. Presently, however, there is an argument to support a regulatory framework that promotes A-AI independence whenever appropriate.

V.C. State and Institutional Supplemental Regulations

This article has focused on FDA's interest in ensuring the safe and effective use of A-AI products, but state and institutional authorities will likely share this interest. Unlike

41 Traczynski & Udalova, *supra* note 18, at 97.

42 Traczynski & Udalova, *supra* note 18, at 93; Yang et al., *supra* note 18, at 192.

other AI technologies, A-AI products are effectively ‘practicing medicine’. Traditionally, state and institutional authorities, not FDA, regulate the practice of medicine by clinicians via state medical boards, physician peer review, and other structures. Therefore, these entities will have an interest in collaborating with FDA to ensure that patients receive quality care. It will also benefit FDA to have these groups involved, as they have traditionally regulated the tools of medical practice rather than the decision makers.

The proposed framework provides a foundation upon which state and institutional authorities can easily impose additional regulations for the oversight of A-AI products. Like nurse practitioners, state and institutional authorities could specify processes for oversight that must be included in any collaborative or supervisory protocol for an A-AI product, tailoring the protocols to the concerns specific to the product. They could also impose certain restrictions on A-AI oversight and require that physicians under their jurisdiction incorporate these restrictions in their collaborative or supervisory protocols. As long as these supplemental policies do not deescalate the oversight level recommended by FDA for a product, they will not conflict with the proposed framework.

Institutional flexibility to add supplemental safeguards and requirements is particularly desirable since institutions will be in the best position to determine whether an A-AI product has been developed and validated using data that are sufficiently representative of the institution’s patient population, provided that these data are available.

V.D. Limiting Off-Label Use

Another benefit of the proposed framework is that it allows FDA to delineate the permitted functions and patient populations for an A-AI product. Without these safeguards in place, states, institutions, and physicians may attempt to expand the scope of an A-AI technology’s indications through quality improvement initiatives, which are not subject to the oversight and regulations that govern clinical research. This includes ethical oversight that ensures safety, equitable selection of research participants, and the presence of clinical equipoise.⁴³ Furthermore, the results of quality improvement initiatives are not submitted to FDA, and in the context of A-AI products, the agency would not be able to make a determination on whether it is safe and effective for a product to perform novel functions and/or treat novel populations, whether the changes to a product’s scope should affect the level of oversight for the technology, and whether the quality improvement initiative was executed and validated using sufficiently representative data.

Due to the high impact that A-AI products will have on the care of patients, it is imperative for investigations of these technologies to be subjected to the highest standards of ethical and regulatory oversight. The proposed framework sets clear parameters for what an A-AI product is authorized to do in the real-world, and relatedly, when expanding the scope of a technology requires submitting the technology to a clinical investigation.

43 Mary Ann Baily et al., *Special Report: The Ethics of Using QI Methods to Improve Health Care Quality and Safety*, 36 HASTINGS CTR. REP. S1, S5 (2006); Ethel Mitty, *Hastings Center Special Report: The Ethics of Using QI Methods to Improve Health Care Quality and Safety*, 22 J. NURS. CARE QUAL. 97 (2007).

V.E. Patient Trust

The proposed framework may foster patient trust in the use of A-AI products. Although the products in collaborative and supervisory protocols may operate somewhat autonomously, patients can take comfort in the fact that a physician is responsible for overseeing the safe and effective operation of these technologies. The requirement for physician oversight and the mechanism for reviewing and revoking a collaborative or supervisory protocol encourage trustworthiness and accountability, two values that Reddy et al.⁴⁴ cite as integral to the governance of AI in health care.

V.F. Other Methods of Oversight

As mentioned in the introduction of this article, the proposed framework may complement the use of data-based methods for real-world performance monitoring. For example, it is inadequate to rely on patients to monitor and report data on real-world performance monitoring given their lack of medical knowledge. The proposed framework identifies a physician who could be responsible for monitoring and reporting data on A-AI products.

The framework could also coexist with other proposals for performance monitoring, such as Babic et al.'s⁴⁵ continuous risk-monitoring approach for learning algorithms. Under this approach, AI technologies would be continuously monitored, including retesting on past cases, simulated performance checks using perturbed data, and adversarial stress tests.

VI. AREAS FOR FUTURE WORK

In addition to developing the criteria FDA uses to determine the oversight level for an A-AI product, other issues need to be addressed before the proposed regulatory approach can be implemented.

VI.A. Preparing Physicians for Oversight

It may be challenging for physicians to design processes for the oversight of an A-AI product without guidance, and it could be helpful to develop templates for collaborative and supervisory protocols. Nurse practitioner regulations provide a starting point for systematically approaching the items that physicians should include in their processes for oversight. For each function of an A-AI product that is toggled 'on' in a collaborative or supervisory protocol, a physician could specify the mechanism;⁴⁶ timing; and frequency of reviewing a decision made by the technology (eg every decision will be reviewed, every 10th decision will be reviewed); and documentation of the review process.⁴⁷

Guidance on these items may be insufficient to ensure that a physician can successfully provide oversight. Additional work is needed to determine when, if ever, a physician needs to receive special training and licensing to create and enter into a collaborative or supervisory protocol with an A-AI product, and if so, whether FDA, state medical boards, or another authority would be responsible for designing and

44 Reddy et al., *supra* note 39, at 494, 495.

45 Boris Babic et al., *Algorithms on Regulatory Lockdown in Medicine*, 366 *Sci.* 1202 (2019).

46 244 Mass. Code Regs. 4.07 (2021).

47 Ark. State Bd. of Nursing, *supra*, note 24.

providing this training and licensing. There could be requirements for special training and licensing for all autonomous products, or requirements could vary by oversight level (eg physicians could deploy ‘collaborative’ level products without special training and licensing, but they could be required to receive special training in order to deploy ‘supervisory’ level products).

Concerns of bias may move regulatory authorities to require special training for physicians who enter into a collaborative or supervisory protocol with an A-AI product. Biased outputs from a product may be challenging to recognize without knowledge of the topic of bias in AI and familiarity with methods for identifying potential biases.

VI.B. Medical Malpractice Liability

It is unclear how the system of oversight in the proposed framework will affect medical malpractice liability, in part because there is limited case law on liability involving medical AI.⁴⁸ For A-AI products that are allowed to practice independently, the absence of physician oversight likely leaves the developers of a product and the health care institution that deploys the product as parties that could potentially face liability.⁴⁹

For A-AI products that require physician oversight, it may be possible for the physician(s) to be held liable for mistakes made by the product. Medical malpractice claims against physicians for the actions of nonphysician providers, such as nurse practitioners, can occur when the physician(s) fail to provide adequate supervision or the nonphysician provider practices outside their approved scope; makes an incorrect diagnosis; performs an inadequate examination of the patient; or misrepresents themselves as a physician to the patient.⁵⁰ Whether and how these standards would be applied to the physician oversight of A-AI products remains to be seen.

VII. CONCLUSION

We predict that the future will bring more A-AI products to health care, and the potential benefits of these products for patients, providers, and health care systems will be tremendous. We also predict that these products will raise novel regulatory issues, including the appropriate mechanisms for real-world oversight. A building tension is that as of now we are attempting to regulate A-AI products as tools of medical practice rather than as medical decision makers and actors. This creates challenges, such as reviewing machine learning algorithms, and undermines the ability of these products to create greater capacity in our health care system while reducing costs.

Our proposed framework for the oversight of A-AI products attempts to resolve this tension by acknowledging that the purpose and utility of these products is to act as decision makers. The framework is built upon lessons from the oversight of nurse practitioners, a system that has been refined and validated over several decades. The oversight of nurse practitioners has generally become less restrictive over time due to the accumulation of evidence suggesting that nurse practitioners are capable of providing safe and effective care and their independent practice leads to benefits for

48 W. Nicholson Price et al., *Potential Liability for Physicians Using Artificial Intelligence*, 322 J. AM. MED. ASSOC. 1765 (2019).

49 *Id.* at 1766.

50 Richard E. Moses & Andrew D. Feld, *Physician Liability for Medical Errors of Nonphysician Clinicians: Nurse Practitioners and Physician Assistants*, 102 AM. J. GASTROENTEROL. 6 (2007).

patients and health care systems.⁵¹ The proposed framework includes three models for the oversight of A-AI products, potentially allowing for a similar evolution as these products become increasingly sophisticated and physicians, patients, and society become increasingly confident in their operation.

CONFLICT OF INTEREST

The authors have no financial, personal, academic, or other conflicts of interest to report.

ACKNOWLEDGMENTS

Carmel Shachar's work was supported in part by a Novo Nordisk Foundation grant for a Collaborative Research Programme (grant agreement number NNF17SA0027784).

51 Emily A. Gadbois et al., *Trends in State Regulation of Nurse Practitioners and Physician Assistants*, 72 *MED. CARE RES. REV.* 200, 210, 211 (2015).