
Editorial

Setting the agenda: an informatics-led policy framework for adaptive CDS

Jeffery Smith

American Medical Informatics Association, Bethesda, Maryland, USA

Corresponding Author: Jeffery Smith, MPP, 4720 Montgomery Ln. 500, Bethesda, MD 20814, USA (jsmith@amia.org)

Received 10 September 2020; Editorial Decision 10 September 2020

On December 5, 2019, the American Medical Informatics Association (AMIA) held its thirteenth annual policy meeting in Washington, DC with the theme “Clinical Decision Support in an Era of Big Data & Machine Learning.”¹

The 2019 Health Informatics Policy Forum was predicated on three basic premises. First, the field of medical informatics credibly lays claim to the intellectual pedigree of clinical decision support (CDS). The founders of the American College of Medical Informatics and AMIA invented CDS, and the wider health informatics community has championed the advancement of CDS for decades. Second, the environment surrounding CDS is evolving rapidly with an exponential growth in health data, combined with growing capacities to store and analyze such data through cloud computing and machine learning (ML). Third, our current governance and policy-making structures are ill-equipped to handle such a dynamic landscape.

The 2019 Policy Forum included submitted content approved by a review committee² and featured keynotes spanning federal officials with specialty knowledge of artificial intelligence (AI) in aeronautics, finance, and health.³ What emerged from this meeting were several components of a policy framework for what the committee deems “Adaptive CDS.” Borrowing from Sim et al.,⁴ we use the term “Adaptive CDS” to describe CDS that can learn and change its performance over time, incorporating new clinical evidence; new data types and data sources; and new methods for interpreting data. Adaptive CDS enables personalized decision support in a way that has not been possible previously because it has the capacity to learn from data and modify recommendations based on those data. Adaptive CDS stands in contrast to “static” CDS, which are those tools that provide the same output (recommendation/guidance) each time the same input is provided without change through use.

We focus on Adaptive CDS because it represents a conceptual use case within a larger ecosystem of potential use cases of AI in

healthcare. By framing our discussion on Adaptive CDS, we hope to engender a practical discussion of policies needed to ensure safe and effective use of AI/ML-driven CDS for patient care and facilitate a wider discussion of policies needed to build trust in the broader use of AI in healthcare.

Many of the review committee members subsequently developed an AMIA Board-approved position paper to establish a policy agenda for the safe, effective use of Adaptive CDS in the US healthcare system—and position AMIA as the organization to lead this agenda’s execution. In this paper, we present a policy framework that spans the design and development, implementation, evaluation, and ongoing maintenance of Adaptive CDS.⁵ This work envisions an extensive policy landscape that includes transparency metrics for Adaptive CDS training datasets, communications standards to provide accurate information about the intended uses of Adaptive CDS, and dedicated actors and protocols to evaluate, test, and monitor Adaptive CDS *in situ*.

The policy framework we describe is not uniquely ours. We rely on decades of scholarship from AMIA members,⁶ including from the paper’s coauthors, and we incorporate ideas discussed at the 2019 Policy Forum. The result is a composite framework encompassing policy concepts and actions necessary for the safe, effective use of Adaptive CDS. If viewed in this way, our policy framework provides a structure for the other papers in this issue of JAMIA to plug in. As a developmental document, our framework describes concepts requiring additional detail and work that will emerge as Adaptive CDS evolves and matures. The 4 accompanying perspectives reflect areas where the informatics community must contribute and should lead.

The AMIA position paper calls for identification of two policy concepts: *transparency metrics* and *communications standards*. Transparency metrics would describe how Adaptive CDS algorithms are trained, including the data acquisition processes (eg, patient co-

hort selection criteria) and preprocessing or “data wrangling” steps that must be clearly documented. Communications standards articulate the components of the Adaptive CDS and describe the intended use(s) and expected user(s), similar to US Food and Drug Administration’s (FDA’s) prescription drug-labeling requirements.⁷

Hernandez-Boussard et al, from Stanford University, describe a reporting standard for AI in healthcare, MINimum Information for Medical AI Reporting, or MINIMAR.⁸ They identify the minimum information necessary to understand an algorithm’s intended predictions, target populations, hidden biases, and ability to generalize to the setting and population in which it is applied. The general principles of the MINIMAR design include features related to (1) the population providing the training data; (2) training data demographics to enable comparison with target population demographics; (3) details about the model architecture to compare to similar models and permit replication; and (4) model evaluation, including optimization and validation, techniques. Alongside CONSORT,⁹ SPIRIT,¹⁰ and TRIPOD,¹¹ MINIMAR adds to the collection of scientific reporting standards necessary to address potential biases and unintended consequences, as well as provide external validation and secondary use of algorithms used in health care. Such reporting standards are an essential component of a policy framework for Adaptive CDS.

The AMIA position paper notes that much of the opportunity for Adaptive CDS lies in the tool’s dynamism and capacity to continuously learn over time, which can only be realized once implemented. This aspect of Adaptive CDS represents the promise and peril of AI/ML-driven applications in healthcare. While evermore personalized care is the goal of continuously learning algorithms, a particularly insidious challenge has emerged: bias. Perspectives from DeCamp and Lindvall¹² and McCradden et al¹³ explore dimensions of bias and offer prescriptions to mitigate its damaging effects by framing bias as a patient safety issue. Ferryman¹⁴ focuses on the FDA’s Precertification Program for AI/ML-driven Software-as-a-Medical Device (SaMD)¹⁵ offering ways to address a distal outcome of algorithmic bias-widening health disparities. These perspectives have important, practical implications for public policy and fill critical gaps in our policy framework.

DeCamp and Lindvall, from the University of Colorado, describe latent bias, akin to latent errors, as bias “waiting to happen” in complex systems.¹² They identify three challenges in dealing with bias in algorithms that have been deemed objectively as “fair,” including (1) adaptive models can become biased over time; (2) AI operates within clinical environments that can be impacted by automation and privilege bias; and (3) bias resulting from a difference in the choice of outcome. They suggest that latent bias be treated as a patient safety issue to be identified and addressed proactively and preferably *ex ante*. They recommend treating biases that emerge over time as adverse events and part of mandatory reporting requirements.

Likewise, McCradden et al, from the Hospital for Sick Children, describe a need to view bias in ML through the lenses of patient safety and quality improvement, with the aim of preventing unintended harms and augmenting the provision of care with respect to equity.¹³ However, their ethical motif, which relies on concepts of nonmaleficence, relevance, accountability, transparency, and equity leads them to a set of requirements for developers to: (1) maintain statistics on the characteristics of the population on which the model was trained and (2) conduct audits of their models consistent with a

focus on continuous quality improvement and with an emphasis on local validation through a prospective, noninterventional silent period.

Ferryman, from New York University, argues that health disparities can be propagated by ML applications in healthcare, and that we must enhance the FDA’s proposed regulation for AI/ML-driven SaMD at both the premarket and postmarket phases.¹⁴ She frames health disparities as a safety issue and recommends that, as part of the premarket assessment, device manufacturers document the representativeness of protected groups in the AI’s training dataset and discuss how different choices made in the design of the model may impact groups already experiencing health disparities. Ferryman also recommends a “Health Equity Review” as part of the FDA’s proposed Real World Monitoring requirements, which would require device manufacturers to examine how the ML tool affects protected groups, whether any new group differences have emerged due to the tool’s use, and describe any activities undertaken to address unfairness or group harms.

These perspectives underscore several points of emerging consensus. First, transparency in how ML-driven decision-making applications are trained is paramount. Without transparency, there can be no accountability. Second, we must develop standards to convey specific attributes of how the model was trained, how it is designed, and how it should operate *in situ* to objectively compare, evaluate, and ensure ongoing maintenance of the algorithm. Third, algorithmic bias must be viewed as a matter of patient safety and quality improvement.

Logically stemming from these points is the practical need to establish agency and oversight—both regulatory and nonregulatory—to manage how these objectives are achieved through consistent systems and controls. With this need in mind, the AMIA position paper calls for creation of new bodies, groups, or departments that govern implementation and use of AI within an institution, as well as a system of oversight across institutions. It also calls for Adaptive CDS Centers of Excellence to develop, test, evaluate, and advance the use of safe, effective ML in practice.

The use of AI in healthcare presents clinicians and patients with opportunities to improve care in unparalleled ways. Equally unparalleled is the opportunity for AMIA and the informatics community to take a position of leadership in ensuring that AI/ML-driven technology is objective, unbiased, safe, and effective. We reiterate that our work is purposefully incomplete. It is the obligation of those who pioneer CDS to continue its stewardship using new and novel technologies. We call on the AMIA membership to join us in creating the necessary policies and standards to fulfill the promise of more patient-centered care through new and novel informatics tools.

CONFLICT OF INTEREST STATEMENT

None declared.

ACKNOWLEDGMENTS

The author would like to acknowledge the contributions of Carolyn Petersen, MS, MBI, FAMIA; Robert R. Freimuth, PhD; Kenneth W. Goodman, PhD; Gretchen Purcell Jackson, MD, PhD, FACS, FACMI, FAMIA; Joseph Kannry, MD; Hongfang Liu, PhD; Subha Madhavan, PhD, FACMI; Dean F. Sittig,

PhD; Adam Wright, PhD as well as the participants of the 2019 Health Informatics Policy Forum who shared their expertise and time.

REFERENCES

- [1]. AMIA 2019 Health Informatics Policy Forum: Clinical Decision Support in the Era of Big Data and Machine Learning. Event Page. <https://www.amia.org/apf2019>.
- [2]. AMIA 2019 Health Informatics Policy Forum Review Committee. <https://www.amia.org/apf2019/steering-committee-members>.
- [3]. AMIA 2019 Health Informatics Policy Forum. Keynotes. <https://www.amia.org/apf2019/keynotes>.
- [4]. Sim I, Gorman B, Greene R, *et al*. Clinical decision support systems for the practice of evidence-based medicine. *J Am Med Inform Assoc* 2001; 8 (6): 527–34.
- [5]. Petersen C, Smith J, Freimuth R, *et al*. Recommendations for the safe, effective use of adaptive CDS in the US Healthcare System. *J Am Med Inform Assoc* forthcoming.
- [6]. Miller RA, Gardner RM, for the American Medical Informatics Association (AMIA), the Computer-based Patient Record Institute (CPRI), the Medical Library Association (MLA), the Association of Academic Health Science Libraries (AAHSL), the American Health Information Management Association (AHIMA), the American Nurses Association. Recommendations for responsible monitoring and regulation of clinical software systems. *J Am Med Inform Assoc* 1997; 4 (6): 442–57.
- [7]. US Food and Drug Administration. Labeling for Human Prescription Drug and Biological Products—Implementing the PLR Content and Format Requirements Final Guidance. 2013. <https://www.fda.gov/media/71836/download> Accessed August 31, 2020
- [8]. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, *et al*. MINIMAR (MINimum Information for Medical AI Reporting): developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc* 2020; 27 (12): 2011–15.
- [9]. Liu X, Faes L, Calvert MJ, Denniston AK. Extension of the CONSORT and SPIRIT statements. *Lancet* 2019; 394 (10205): 1225.
- [10]. Chan AW, Tetzlaff JM, Altman DG, *et al*. SPIRIT 2013 Statement: defining standard protocol items for clinical trials. *Rev Panam Salud Publica* 2015; 38 (6): 506–14.
- [11]. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; 162 (1): 55–63.
- [12]. DeCamp M, Lindvall C. Latent bias and the implementation of artificial intelligence in medicine. *J Am Med Inform Assoc* 2020; 27 (12): 2020–23.
- [13]. McCradden MD, Joshi S, Anderson JA, *et al*. Patient safety and quality improvement: ethical principles for a regulatory approach to bias in healthcare machine learning. *J Am Med Inform Assoc* 2020; 27 (12): 2024–27.
- [14]. Ferryman K. Addressing health disparities in the Food and Drug Administration’s artificial intelligence and machine learning regulatory framework. *J Am Med Inform Assoc* 2020; 27 (12): 2016–19.
- [15]. Food & Drug Administration. Software Precertification Program: Working Model v1.0. 2019. <https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/UCM629276.pdf>.