AMIA Position Paper

Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: an AMIA position paper

Carolyn Petersen [1], Jeffery Smith[2], Robert R. Freimuth[3], Kenneth W. Goodman [4], Gretchen Purcell Jackson[5,6], Joseph Kannry[7], Hongfang Liu[8], Subha Madhavan[9], Dean F. Sittig [10], and Adam Wright[11]

[1]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota, USA, [2]The Office of the National Coordinator for Health Information Technology, Washington, DC, USA, [3]Division of Digital Health Sciences, Center for Individualized Medicine, Mayo Clinic, Rochester, Minnesota, USA, [4]Institute for Bioethics and Health Policy, University of Miami Miller School of Medicine, Miami, Florida, USA, [5]IBM Watson Health, Cambridge, Massachusetts, USA, [6]Department of Pediatric Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, USA, [7]Mount Sinai Health System, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [8]Division of Digital Health Sciences, Mayo Clinic, Rochester, Minnesota, USA, [9]Department of Oncology, Georgetown Lombardi Comprehensive Cancer Center, Innovation Center for Biomedical Informatics, Georgetown University Medical Center, Washington, DC, USA, [10]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, UT-Memorial Hermann Center for Healthcare Quality & Safety, Houston, Texas, USA, [11]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Corresponding Author: Carolyn Petersen, MS, MBI, FAMIA, 200 First St. SW, Rochester, MN 55905, USA; petersen.carolyn@mayo.edu

ABSTRACT

The development and implementation of clinical decision support (CDS) that trains itself and adapts its algorithms based on new data—here referred to as Adaptive CDS—present unique challenges and considerations. Although Adaptive CDS represents an expected progression from earlier work, the activities needed to appropriately manage and support the establishment and evolution of Adaptive CDS require new, coordinated initiatives and oversight that do not currently exist. In this AMIA position paper, the authors describe current and emerging challenges to the safe use of Adaptive CDS and lay out recommendations for the effective management and monitoring of Adaptive CDS.

Key words: artificial intelligence, machine learning, software as a medical device, clinical decision support, health policy

INTRODUCTION

Since the passage of the Health Information Technology for Economic and Clinical Health Act, the landscape of health informatics tools supporting these transformations is evolving in dramatic ways. The quantity and types of health data are growing at exponential rates as electronic health records (EHRs) achieve ubiquity in clinical care settings,[1,2] new molecular/omics data are used for precision medicine,[3] and the use of consumer applications grows outside traditional care settings.[4,5] Growing capacities to store and analyze such data through cloud computing and machine learning and the development of standards to enable the exchange and use of those data by other systems support this rapid expansion of health data.

Clinical decision support (CDS) represents one of the most important applications of computing to patient care and continues to advance as does the widespread deployment of EHRs and the professionalization of clinical informatics. Contemporary CDS tools are

designed to provide patient-specific, timely, and appropriate recommendations, including risk assessment (developing disease in the future), diagnosis (determining the presence or absence of disease at the current time), prognosis (forecasting the likely course of a disease), therapeutics (predicting treatment response), medication error control, and much more.

Recent trends in health IT have created new opportunities to develop and deploy decision support systems that adapt to dynamic and growing bodies of knowledge and data, and are 1) increasingly driven by artificial intelligence (AI) and machine learning (ML) algorithms; 2) available as substitutable knowledge resources;[6] and 3) used as tools to diagnose diseases and conditions in Software-as-a-Medical Device (SaMD) products.[7] These trends offer significant promise for reducing alert fatigue, improving cognitive load, and delivering better evidence to point-of-care decisions in the form of Adaptive CDS.

In this AMIA position paper we use the term "Adaptive CDS" to describe CDS that can learn and change performance over time, incorporating new clinical evidence, new data types and data sources, and new methods for interpreting data. Others have applied the term "Adaptive CDS" to mean post-implementation surveillance of CDS,[8] but here we frame Adaptive CDS as an AI use case. 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 that provide the same output (recommendation/guidance) each time the same input is provided, resulting in output that does not change with use. An example of a static CDS is the Atherosclerotic Cardiovascular Disease (ASCVD) risk score event, which predicts the risk of myocardial infarction or stroke over the following 10 years.[9] Adaptive CDS changes its output using a defined learning process that is based on new knowledge or data. In the above example, we envision that an adaptive ASCVD risk score would be refined based on data from an institution where it is deployed. Such functionality could be applied to a wide range of tasks, including risk assessment, prognosis tools, treatment planning, diagnostic support, clinical trial recruitment, and treatment support, among others.

Two categories of Adaptive CDS warrant distinction for purposes of establishing public policy. First, we refer to Adaptive CDS that is sold to customers for use in a healthcare setting as Marketed ACDS. Second, we refer to Adaptive CDS that is developed in-house by healthcare systems and not marketed or sold to others as Self-Developed ACDS. Marketed ACDS is subject to FDA oversight per the 21st Century Cures Act[10] and related FDA interpretation.[11] Self-Developed ACDS is very likely unregulated by any federal entity and is already used routinely without oversight by any authoritative body—public, private, or nonprofit.

Debates about the scope and force of oversight for the safety and effectiveness of CDS have tended to emphasize legal regulation, of which little exists, and institutional governance, which frequently is wholly lacking. Organizational leaders have recognized content creation, analytics and reporting, and governance and management as critical components in the development of CDS, but achieving all 3 in sufficient depth remains challenging for organizations.[12]

These recommendations (see Table 1) articulate policies needed to ensure the safe and effective use of both Marketed and Self-Developed ACDS in clinical settings. We suggest that both kinds of Adaptive CDS should be underpinned by robust testing and transparency metrics for training datasets. Additionally, we argue that consistency is needed to communicate significant aspects or elements of the Adaptive CDS. Finally, we discuss the urgent need to 1) artic-

ulate which aspects of FDA's safety and efficacy regulatory controls may pertain to Self-Developed ACDS; 2) identify stakeholders who may be best positioned to execute nonregulatory controls; and 3) develop a strategy to implement pre- and postmarket oversight for Self-Developed ACDS.

During the 2019 AMIA Health Informatics Policy Forum, "Clinical Decision Support in the Era of Big Data and Machine Learning,"[13] AMIA members and collaborators discussed opportunities and challenges of Adaptive CDS given the growing volume, variety, velocity, and veracity of data, the increasing diversity of data sources, the advancing computational power of artificial intelligence, and growing societal expectations of technology as a promoter of health equity despite inherent and increasingly recognizable biases. Informed by discussions at the Policy Forum, a group of forum planners and attendees identified key learnings, forming the foundation for AMIA's vision for a policy and research agenda that promotes the safe, ethical, and effective design, development, and use of Adaptive CDS in clinical settings. We focus on the use case of Adaptive CDS because it represents a wide range of potential tools and applications—some examples of which exist today, but many of which do not; because Adaptive CDS represents a conceptual use case within a larger ecosystem of potential use cases of AI in healthcare; and because Adaptive CDS represents an area (decision support) in which the informatics community has strong intellectual roots and expertise.

The current policy and oversight landscape for Adaptive CDS

In the United States, several federal agencies within the Department of Health and Human Services (HHS) have purview over the safety, use, and functionality of health IT, including CDS. For example, the HHS Office of the National Coordinator for Health IT (ONC) certifies and regularly updates CDS functional requirements, as well as the underlying standards for interoperability.[14] Reimbursement requirements established by the Centers for Medicare and Medicaid Services (CMS) encourage continued and widespread adoption of CDS among providers participating in those programs.[15,16] The Food and Drug Administration (FDA) oversees medical device products and the data systems that support such products (eg, Medical Device Data Systems).[17,18] Strict lines of jurisdiction exist across CMS, ONC, and FDA regarding the practice of medicine, leading to significant gaps in oversight for emerging technology. The use of Adaptive CDS in clinical practice presents a unique challenge to these jurisdictional boundaries and has already led to a new approach at FDA known as the Pre-Certification (Pre-Cert) Program. This approach shifts the emphasis of regulation from the product's technical functionality to the firm that developed the device, relying on an assessment of whether the firm's software development practices "provide reasonable assurance of safety and effectiveness of the organization's [software] products."[19] Although the object of this specific regulatory approach by the FDA is SaMD, many kinds of marketed CDS—especially AI/ML-driven CDS—will be treated in a similar manner.[11]

NECESSARY STEPS TO ENSURE SAFE AND EFFECTIVE DESIGN AND DEVELOPMENT OF ADAPTIVE CDS

Significant gaps between the design and development of CDS and their integration into clinical workflow exist.[20] This circumstance has hindered or frustrated adoption when CDS systems are intro-

Table 1. AMIA position paper recommendations

1. **FDA's Precertification Program and Total Product Life Cycle regulatory approach should include transparency metrics and communication standards, developed by multidisciplinary and multisector stakeholders.**
   a. Transparency metrics include:
      i. Metrics regarding algorithm training datasets, including semantic and provenance measures, and detailed descriptions of the data acquisition process, such as selection criteria of cohorts
      ii. Additional measures that communicate the nature and prevalence of attributes likely to influence how a model may perform on new data
      iii. Descriptions of preprocessing or "data wrangling" steps made available with the model that elucidate the representativeness (or lack thereof, aka "missingness") of the training data, to mitigate bias
      iv. Explicit definitions and assessment of inter-annotator agreement
      v. Metrics for the process of generating and labeling datasets
      vi. Criteria and/or events guiding the scheduling of evaluation or testing

   b. Communication standards include:
      i. A range of information pertaining to the expected performance and limitations of the ACDS including:
         1. Expected users and uses
         2. Indications and usage
         3. Contraindications
         4. Warnings and precautions
         5. Interactions and dependencies
         6. Adverse outcomes
      ii. Reports (publicly available, searchable online, current) describing the results of tests/evaluations and fine tuning

2. **FDA should require manufacturers of Adaptive CDS to develop and present Adaptive CDS Assurance Case(s) for regulatory approval through its Precertification Program and fulfillment of its real-world performance requirements.**
   An Adaptive CDS Assurance Case helps establish institutional governance and identify performance indicators to transition from the evaluation phase to the deployment phase. An Adaptive CDS Assurance Case includes:
   a. Claims related to the stated clinical utility for the Adaptive CDS, using the communications standards and describing which patient/populations are relevant
   b. Transparency metrics describing how the Adaptive CDS was trained and tested
   c. Expected deployment challenges
   d. Claims related to ongoing maintenance of the algorithm and claims related to when periodic evaluation of the algorithm should take place

3. **Dedicated funding from the NIH, FDA, CDC, ONC, and CMS should establish Adaptive CDS Centers of Excellence to develop, test, evaluate, and advance the use of safe, effective ACDS in practice.**
   Such ACDS Centers of Excellence should be charged with:
   a. Advancing post-market surveillance strategies
   b. Developing ongoing guidance for how best to deploy and maintain Adaptive CDS for patient care
   c. Serving as clearinghouses for new knowledge regarding potential negative impacts and biases resulting from the use of Adaptive CDS

4. **Non-regulatory oversight bodies and programs should be established for Self-Developed Adaptive CDS used in clinical settings.**
   Self-Developed ACDS should be managed by a combination of local groups, known as Adaptive CDS Review Groups, and a national Accreditation Commission(s):
   a. Adaptive CDS Review Groups are local or institutional bodies, groups, or departments that are well-versed and empowered to require or negotiate aspects of ACDS implementation.
   b. Adaptive CDS Accreditation Commission is a body responsible for ensuring that ACDS Review Groups have the requisite skills and resources to evaluate Assurance Cases and perform on-site testing of Self-Developed ACDS once deployed. There could be a single national group, or a small number of groups established as an Accreditation Commission.
      i. ACDS Centers of Excellence would contribute, advise, or support such a body in their work.

5. **Medical education curricula should be modified to include analysis and evaluation of Adaptive CDS.**
   Medical education and knowledge management must be updated to provide students and experienced clinicians with exposure to Adaptive CDS across all training phases and be maintained as part of professional development for practicing clinicians.

duced at the point of care, creating unanticipated negative effects on clerical burden, care team communication, and shared understanding between providers and patients.[21] Although CDS may support patient–provider relationships when providers share screen viewing with the patient, CDS also may draw providers' attention away from patients or create additional cognitive burden for clinicians.[22] Adaptive CDS will exacerbate these and related challenges without increased emphasis on design and development protocols. Transparency in design and development ought to be the preeminent value in

software design because without transparency, accountability and responsibility are rendered meaningless.

Premarket oversight for Adaptive CDS

In developing its Pre-Cert Program and associated Total Product Life Cycle (TPLC) regulatory approach, FDA sought to shift scrutiny from premarket data collection and submission to postmarket performance monitoring. Rightly, this shift has occurred because
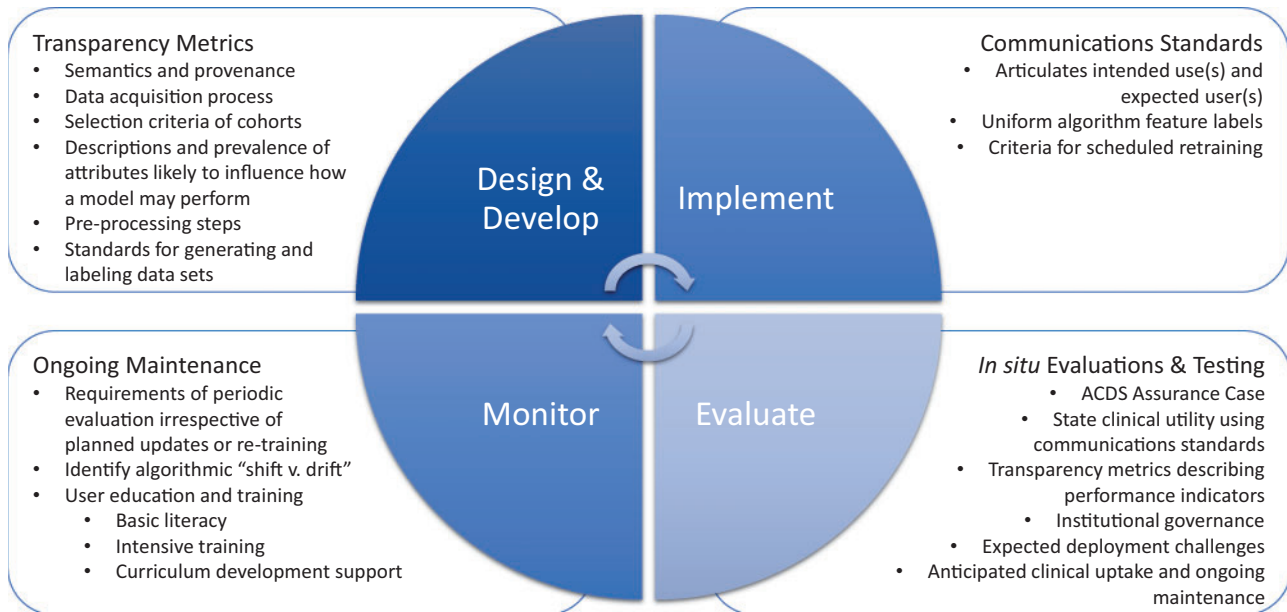
Transparency Metrics
- Semantics and provenance
- Data acquisition process
- Selection criteria of cohorts
- Descriptions and prevalence of attributes likely to influence how a model may perform
- Pre-processing steps
- Standards for generating and labeling data sets

Communications Standards
- Articulates intended use(s) and expected user(s)
- Uniform algorithm feature labels
- Criteria for scheduled retraining

Design & Develop

Implement

Monitor

Evaluate

Ongoing Maintenance
- Requirements of periodic evaluation irrespective of planned updates or re-training
- Identify algorithmic "shift v. drift"
- User education and training
  - Basic literacy
  - Intensive training
  - Curriculum development support

***In situ* Evaluations & Testing**
- ACDS Assurance Case
- State clinical utility using communications standards
- Transparency metrics describing performance indicators
- Institutional governance
- Expected deployment challenges
- Anticipated clinical uptake and ongoing maintenance

Figure 1. Policy recommendations for all stages of Adaptive CDS (ACDS)—design and development, implementation, evaluation, and ongoing monitoring—require further development to ensure safe and effective ACDS. A concerted multistakeholder effort to identify key transparency metrics for training datasets and communications standards for AI-driven applications in healthcare is needed to understand how bias can corrupt AI-driven decision support and identify ways to mitigate such bias. Additionally, policies that standardize *in situ* testing and evaluation, as well as ongoing maintenance, of ACDS should be established.

much of the opportunity for Adaptive CDS lies in the tool's dynamism and capacity to continuously learn, which can only be realized once implemented. However, there is also a pressing need to advance premarket transparency metrics and communications standards related to how Adaptive CDS was developed and tested. Figure 1 depicts the policy recommendations for all stages of ACDS, which are described in greater detail below.

Adaptive CDS design and development transparency metrics

Transparency metrics are needed for training datasets used to design, develop, and implement Adaptive CDS. Specifically, the semantics and provenance of these datasets must be unambiguously communicated and available for validation prior to deployment. Additionally, detailed descriptions of the data acquisition process, selection criteria of cohorts, and descriptions and prevalence of attributes likely to influence how a model may perform on new data are needed.

Apart from these attributes of training datasets, preprocessing or "data wrangling" steps must be clearly documented and made available with the model to help explain the representativeness (or lack thereof) of the training data and to help identify areas where bias may exist. For example, how one deals with missing values ("missingness" in statistics parlance) or sample imbalances must not be overlooked or omitted. These choices, and the preprocessing of data before learning a model is apace, mark the venues where bias can be introduced. Meanwhile, comprehensive and transparent annotation requires explicit definitions and the assessment of inter-annotator agreement, which will help prevent bias introduced by individual preferences. The adoption of standards for scientific rigor in software engineering, perhaps especially in the process of generating and labeling datasets for specific AI tasks, is necessary to ensure the validity of underlying models.

When evaluating learning algorithms in Adaptive CDS, it is crucial to have simple and obvious baselines on the same dataset. Clarity

and transparency foster honest comparisons of "black box" solutions derived from deep learning and explainable AI models, such as decision trees. Aside from the need to develop these transparency metrics, methodological issues in continuous evaluation (or testing) of an AI model as underlying data change and evolve may require retraining of the original models. For example, developers could retrain every week a model for predicting 24-hour mortality as new data become available. Developers should explicitly clarify and justify the criteria and/or events guiding this scheduling. Interesting and potentially deleterious challenges, as well as social inequities, may arise if a model that is somehow "approved" can be retrained the next week and behave differently. Articulating foundational assumptions and the basis for training recommendations can aid in monitoring data model performance to avoid adverse effects of continuously evolving models. The Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) work, which is designed to improve reporting of predictive model studies, may offer a basis for transparent examination, implementation, and assessment of predictive models used in Adaptive CDS.[23,24]

Adaptive CDS communications standards

In addition to transparency metrics for dataset attributes, premarket requirements for Adaptive CDS should include standards for communicating specific aspects of the technology, including expected performance parameters and limitations of the Adaptive CDS. Developers should be able to, and should, address a range of questions regarding the intended use and expected users of Adaptive CDS in a consistent manner. FDA prescription drug labeling regulations have been suggested for AI/ML-driven SaMD, and there are relevant correlates. According to the agency, the 2006 Physician Labeling Rule "was designed to make information in prescription drug labeling easier for healthcare practitioners to access, read, and use to facilitate practitioners' use of labeling to make prescribing decisions."[25] It included updates to how prescription drugs communi-

cated concepts, such as indications and usage; contraindications; warnings and precautions; interactions; adverse reactions; and use in specific populations. These communication concepts are relevant to the intended use and expected users of Adaptive CDS. ACDS may not be useful in specific clinical settings or for specific clinical purposes, so such standards for communicating how a clinician should apply Adaptive CDS are needed.

Finally, Marketed ACDS should be tested and optimized before approval for use by FDA. Just as policy can create transparency metrics for how the Marketed ACDS was developed and tested, and communication standards help users determine which ACDS is best for their patients/populations, there should also be a reporting requirement that describes the results of such tests and fine tuning. This information on premarket FDA approval should be publicly available, searchable online, and current. The FDA's database of searchable de novo and 510(k) devices should be leveraged for such reports and updated to improve functionality.

POLICIES TO ENSURE SAFE AND EFFECTIVE DEPLOYMENT AND USE OF ADAPTIVE CDS

As healthcare institutions escalate investment in AI, they face enormous pressure for rapid results. In particular, the lowered technical hurdle for creating Adaptive CDS further contributes to the tendency to deploy such tools in routine clinical care without an adequate understanding of the tools' validity and potential influence.[26] Users have given limited attention to significant risks, both in terms of wasted resources and poor decisions backed by inaccurate prediction.[27,28] Some healthcare settings lack the capacity to critically evaluate the products they purchase or develop in-house, which may leave clinicians open to software malfunction and patient harm, and patients open to ineffective and potential damaging outcomes.

Previously published work recognizes the need for a defined standard that supports understanding of intended predictions, target populations, hidden biases, and other information needed to assess AI-based CDS.[29] Latent biases—biases "waiting to happen" that have not been purposefully avoided or eliminated—pose an additional, and no less important threat, and require their own approaches for identification and management.[30] The potential for AI-generated bias to increase health disparities is a real possibility, and one of particular concern given the imperative to find effective treatments for COVID-19,[31] whose effects are experienced disproportionately by African Americans and Latinos.[32] Evaluation and ongoing maintenance of Adaptive CDS, combined with workforce education and knowledge management of such applications, can help to ensure the safe and effective deployment and use of Adaptive CDS in clinical settings.

Evaluation

Organizations should evaluate health information systems, perhaps especially in hospitals, in the contexts of systems in which they will be used.[33] Efficient and "field-tested" measures are indispensable to the successful utilization of AI in healthcare, and institutional policies and any government regulation must require them. Ideally, appropriate oversight will help patients, clinicians, researchers, payers, and institutions better understand and assess the quality characteristics of Adaptive CDS. These means facilitate establishment of evidence-based trust before CDS systems are adopted into practice and workflow.

Institutional governance and performance indicators are necessary for the ongoing maintenance and review of Adaptive CDS once deployed. One mechanism to facilitate institutional governance and identify performance indicators is to transition from the evaluation phase to the deployment phase with an Adaptive CDS Assurance Case. Borrowed from the ISO/IEC concept,[34] an assurance case is 1) a top-level claim for a property of a system or product (or set of claims for a system or product); 2) a systematic argumentation regarding this claim; and 3) the evidence and explicit assumptions that underlie this argument. Arguing through multiple levels of subordinate claims, this structured argumentation connects the top-level claim to underlying evidence and assumptions. For the purposes of Adaptive CDS, these assurance cases could include claims related to:

- The stated clinical utility for the Adaptive CDS, using the communications standards describing which patients/populations are relevant
- Transparency metrics describing how the Adaptive CDS was developed and tested
- The feasibility of successful implementation given available data
- Expected deployment challenges
- Anticipated clinical uptake and associated metrics
- Ongoing maintenance

In addition to this structured argumentation, ethical issues require explicit evaluation of performance. Well-documented decision tools, especially risk stratification algorithms, have been found to perpetuate racial bias, sometimes affecting millions of patients.[35] Likewise, concerns are growing among patients and their families who are often neither informed about nor asked to consent to the use of tools that help predict hospitalizations, treatment complications, or risk of readmission.[36] The increased reliance on algorithms to guide health decisions demands regular, ongoing, and structural review to determine whether racial, socioeconomic, age, or other biases occur during use. An approach grounded in patient safety and quality improvement principles offers 1 way to identify and analyze how bias affects machine learning and thereby mitigate harm and facilitate transparency and accountability.[37]

Ongoing maintenance and review

A 2019 white paper articulated FDA's view about how to manage changes to continuously learning SaMD through a TPLC approach to regulation.[38] The FDA's Modification Framework instigated discussion about real-world performance monitoring for SaMD, which would be similar in nature to Adaptive CDS. The FDA posited that "The predetermined change control plan would include the types of anticipated modifications—SaMD Prespecifications—based on the retraining and model update strategy, and the associated methodology—Algorithm Change Protocol—being used to implement those changes in a controlled manner that manages risks to patients."[38] "Algorithm Change Protocol" as used by FDA refers to an algorithm's ability to change over time as a result of input—what, in this position paper, is referred to ACDS. In comments submitted to FDA, AMIA made several recommendations worth consideration as part of Adaptive CDS postdeployment maintenance and review.[39] AMIA recommended the Modification Framework include requirements of periodic evaluation irrespective of planned updates or retraining. AMIA further recommended that FDA seek additional feedback to understand a basis for determining when periodic evaluation should occur. We anticipate that notifications to FDA about changes to software could be deterministic—triggered when a

threshold of data processing or algorithmic adaptations have occurred and/or the lapse of a specific time interval (eg, a year). As akin to Genetic Shift, this could be considered Algorithm Shift.

However, conditions may occur in which the AI/ML-based SaMD's behavior changes due to real-time changes in its inputs/change protocol/outputs, via Algorithm Drift. In this scenario, there needs to be some regulatory requirement that, even when the target population and indication do not change, incremental change (Algorithm Drift) in the SaMD is compared with a static historical control. New standard drifts from the historical control by some predetermined amount (eg, a *P* value or a percentage, depending on the output and input) must trigger an FDA review. An annual report to FDA indicating whether the AI/ML-based SaMD has changed as the result of Shift or Drift is an important component to transparency and real-world performance monitoring. We also note that a lack of change may be cause for concern among continuously learning SaMD. Other work offers analysis of FDA's proposed regulatory framework with regard to health disparities and recommends premarket and postmarket practices to advance health equity.[40]

Workforce education and knowledge management

Evaluation and monitoring must be paired with adequate user education and deliberate knowledge management. There is a great need for comprehensive, targeted education to shape Adaptive CDS applications and guide users. AMIA has historically advocated for 3 levels of education and training:[41]

- Basic "informatics literacy" for all health professionals that goes beyond computer or health IT literacy. Literacy in informatics should become part of medical education, biomedical research, and public health training to give clinicians the skills needed to collect and analyze information and apply it in their practice.
- Intensive applied informatics training to improve leadership and expertise in applying informatics principles to the collection and analysis of information and its application to healthcare problems. This level of training will ensure a supply of qualified professionals for the emerging roles of chief medical information officers, chief nursing information officers, chief clinical informatics officers, chief research officers, and similar roles.
- Support for education professionals who will advance the science and train the next generation of informatics professionals in this developing and dynamic field of study.

Similarly, professional development of students and experienced clinicians in the health professions would benefit from training in evaluation of CDS and Adaptive CDS. Just as we now teach students how to read the literature, assess the quality of clinical studies, and search PubMed and CINAHL, so should we introduce analysis and evaluation of Adaptive CDS at this earliest stage of training. This exposure should continue through all training phases, and just-in-time education must be provided for experienced clinicians. Addressing knowledge management would provide a way to curate and organize the content areas for review by attending clinicians. Education about knowledge management is sorely lacking, and if we are to make the most effective use of decision support it will be best realized as a tool that runs brightly through all aspects of clinical practice.[42]

At the same time, we recognize that patients act as developers and users of algorithm-based technologies, such as the Nightscout Project and the Open Artificial Pancreas System.[43,44] To most safely and effectively develop and use such technologies, patients must be able to avail themselves of professional-quality training. We there-fore support the inclusion of patients in ACDS-focused educational programming to ensure mutual benefit to both commercial and non-commercial developers.

A POLICY FRAMEWORK FOR SELF-DEVELOPED ACDS

Given FDA's regulatory purview and competencies, we articulate here a policy framework for Self-Developed ACDS. This framework describes a system that encourages institutional governance and is built upon a network of nonregulatory oversight for the various categories and enumerable uses of Adaptive CDS that are developed by healthcare organizations for use internally. While we do not call for FDA regulation of Self-Developed ACDS, we do encourage the support and use of this policy framework by the FDA for Marketed ACDS. Consistency in how both kinds of ACDS are overseen will be beneficial, especially as FDA's envisioned real-world performance analytics framework develops.[45]

This framework incorporates ethical issues identified over more than 3 decades. An important initial analysis highlighted the importance of identifying appropriate uses and users of decision-support tools; signaled the need for user education; and introduced the idea that failure to use a computer program to improve care might be as blameworthy as its inappropriate use.[46] Those insights have expanded to include assessment of the necessity of CDS evaluation and standards, the role of software engineering, and the importance of public trust.[47]

Several professions (eg, nursing, medicine, and pharmacy) comprising or contributing to biomedical informatics are closely governed by state and federal laws; other professions, such as computer science and laboratory science, are more loosely governed by institutional and professional policies.[48] The experience of nearly a half-century indicates that the hybridized science of biomedical informatics requires some measure of transparency and oversight or governance. However, the degree of oversight, the public/private nature of oversight, and the mechanisms implementing such oversight have proven challenging. The safe and effective use of Adaptive CDS requires, at a minimum, compelling guidance by professional organizations, in addition to policies and support for policy development at institutions that use ACDS systems. One thoughtful and workable approach would be to require creation of internal bodies, groups, or departments[49] that act as a "mechanism for ensuring institutional oversight and responsibility [. . . and] give appropriate weight to competing ethical concerns in the context of internal research for quality control, outcomes monitoring, and so on."[49] These Adaptive CDS Review Groups would need to be well-versed and then empowered to require or negotiate aspects of ACDS implementation. At an institutional level, policies to govern such implementation and use would fall under the remit of such internal bodies, perhaps in consultation with other IT units and institutional ethics committees. A more contemporary, yet similar, suggestion is the formation of a cross-disciplinary, dedicated clinical AI department, which would have purview of the use of AI in healthcare delivery.[50]

We envision that such Adaptive CDS Review Groups would themselves be overseen by an accreditation body similar to the Commission on Accreditation for Health Informatics and Information Management.[51] Such an accreditation body would both ensure that the Adaptive CDS Review Groups are appropriately staffed, skilled, and resourced to evaluate Adaptive CDS Assurance Cases submitted by organizations wishing to deploy Self-Developed ACDS and per-

form on-site testing of Self-Developed ACDS once deployed. This *in situ* testing would ensure that the Assurance Case underlying the use of the ACDS is accurate and that underlying performance indicators are being met. Such an accreditation commission would also ensure that ACDS Review Groups display a requisite level of transparent accountability to the patients that healthcare organizations purport to serve.

Additionally, we recommend that ACDS Centers of Excellence be established to develop, test, evaluate, and advance the use of safe, effective ACDS in practice. With dedicated funding from the NIH, FDA, CDC, ONC, and CMS, ACDS Centers of Excellence would be charged with advancing postmarket surveillance strategies and develop ongoing guidance for how best to deploy and maintain Adaptive CDS for optimal patient care. Centers of Excellence would also serve as clearinghouses for new knowledge regarding potential negative impacts and biases resulting from the use of Adaptive CDS. Centers of Excellence could serve as part of the accreditation commission or support such a body in their work.

CONCLUSION

The use of AI in healthcare presents clinicians and patients with opportunities to improve care in unparalleled ways. As an evolving application of AI in healthcare, Adaptive CDS engenders a practical discussion of policies needed to ensure the safe, effective use of AI-driven CDS for patient care and facilitates a wider discussion of policies needed to build trust in the broader use of AI in healthcare. It is time to redouble our efforts to safely harness the growing volume, variety, velocity, and veracity of data and knowledge, the increasing diversity of data sources, and the advancing computational power of machine learning.

New and flexible oversight structures that evolve with the healthcare ecosystem are needed, and these oversight mechanisms should be distributed across institutions and organizational actors. Furthermore, new organizational competencies are needed to evaluate and monitor Adaptive CDS in patient settings. Although the FDA is developing and testing policies for Marketed ACDS, numerous algorithm-driven applications are self-developed without even cursory guidance or oversight. We believe in the promise of Adaptive CDS and support the use of public policy and federal action to ensure safe and effective deployment.

FUNDING

AUTHOR CONTRIBUTIONS

All the authors made substantial contributions to the conception of the topic and the ideas in the manuscript, and participated in the writing and editing of the position statement. JS developed the figure and table. CP and JS revised the position statement for style and clarity after receiving approval of the position statement from the AMIA Board of Directors.

DATA AVAILABILITY STATEMENT

No new data were generated or analyzed in support of this research.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Office of the National Coordinator for Health Information Technology. Office-based physician electronic health record adoption. 2019. https://dashboard.healthit.gov/quickstats/pages/physician-ehr-adoption-trends.phpAccessed June 11, 2020.
2. Office of the National Coordinator for Health Information Technology. Percent of hospitals, by type, that possess certified health IT. 2018. https://dashboard.healthit.gov/quickstats/pages/certified-electronic-health-record-technology-in-hospitals.php Accessed June 11, 2020.
3. Office of the National Coordinator for Health Information Technology. Sync for Genes: Enabling Clinical Genomics for Precision Medicine via HL7 Fast Healthcare Interoperability Resources. 2017. https://www.healthit.gov/sites/default/files/sync_for_genes_report_november_2017.pdf Accessed August 29, 2020.
4. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019; 381 (20): 1909–17.
5. Tully J, Dameff C, Longhurst CA. Wave of wearables: clinical management of patients and the future of connected medicine. *Clin Lab Med* 2020; 40 (1): 69–82.
6. Richesson RL, Bray BE, Dymek C, et al. Summary of second annual MCBK public meeting: mobilizing computable biomedical knowledge – a movement to accelerate translation of knowledge into action. *Learn Health Sys* 2020; 4 (2): e10222.
7. Food and Drug Administration. De Novo classification request for IDx-DR. https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180001.pdf Accessed June 11, 2020.
8. Khan S, Richardson S, Liu A, et al. Improving provider adoption with adaptive clinical decision support surveillance: an observational study. *JMIR Hum Factors* 2019; 6 (1): e10245.
9. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2014; 63 (25): 2935–59.
10. Section 3060(a) of the 21st Century Cures Act of 2016. PL114–255. 2016. https://www.govinfo.gov/content/pkg/PLAW-114publ255/pdf/PLAW-114publ255.pdf Accessed June 11, 2020.
11. Food & Drug Administration. Clinical Decision Support Software. Draft Guidance for Industry and Food and Drug Administration Staff. September 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software Accessed June 11, 2020.
12. Orenstein EW, Muthu N, Weitkamp AO, et al. Towards a maturity model for clinical decision support operations. *Appl Clin Inform* 2019; 10 (05): 810–9.
13. American Medical Informatics Association. *2019 AMIA Health Informatics Policy Forum*. Washington, DC: American Medical Informatics Association; 2019. https://www.amia.org/apf2019 Accessed June 11, 2020.
14. Office of the National Coordinator for Health IT. Federal Health IT Certification Program. 45 CFR Part 170. 2019. https://www.healthit.gov/topic/certification-ehrs/about-onc-health-it-certification-program Accessed June 11, 2020.
15. Centers for Medicare & Medicaid Services. Promoting Interoperability Program.2020. https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms Accessed June 11, 2020.
16. Centers for Medicare & Medicaid Services. Appropriate Use Criteria Program. 2020. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Appropriate-Use-Criteria-Program Accessed June 11, 2020.
17. Food & Drug Administration. Medical Device Regulations. 21 CFR Part 800. 2018. https://www.fda.gov/medical-devices/device-advice-compre-

hensive-regulatory-assistance/overview-device-regulation Accessed June 11, 2020.

18. Food & Drug Administration. Guidance on Medical Device Data Systems. September 2019. https://www.fda.gov/medical-devices/general-hospital-devices-and-supplies/medical-device-data-systems Accessed June 11, 2020.

19. Food & Drug Administration. Software Precertification Program: Working Model – Version 1.0 – January 2019. 4.2 Excellence Appraisal Elements. https://www.fda.gov/media/119722/download Accessed June 11, 2020.

20. Wright A, Ash JS, Aaron S, et al. Best practices for preventing malfunctions in rule-based clinical decision support alerts and reminders: results of a Delphi study. *Int J Med Inform* 2018; 118: 78–85.

21. Teich JM, Osheroff JA, Pifer EA, et al. Clinical decision support in electronic prescribing: recommendations and an action plan: report of the joint clinical decision support workgroup. *J Am Med Inform Assoc* 2005; 12 (4): 365–76.

22. Richardson S, Feldstein D, McGinn T, et al. Live usability testing of two complex clinical decision support tools: observational study. *JMIR Hum Factors* 2019; 6 (2): e12471.

23. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD group. *Circulation* 2015; 131 (2): 211–9.

24. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162 (1): W1–73.

25. Food & Drug Administration. Labeling for Human Prescription Drug and Biological Products – Implementing the PLR Content and Format Requirements Final Guidance. February 2013. https://www.fda.gov/media/71836/download Accessed June 11, 2020.

26. Ross C. Hospitals are using AI to predict the decline of Covid-19 patients—before knowing it works. STAT. 24 April 2020. https://www.statnews.com/2020/04/24/coronavirus-hospitals-use-ai-to-predict-patient-decline-before-knowing-it-works/ Accessed June 11, 2020.

27. Wolff J, Pauling J, Keck A, et al. The economic impact of artificial intelligence in health care: systematic review. *J Med Internet Res* 2020; 22 (2): e16866.

28. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 2020; 368: m689.

29. 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): 2011–5.

30. DeCamp M, Lindvall C. Latent bias and the implementation of artificial intelligence in medicine. *J Am Med Infor Assoc* 2020; 27 (12): 2020–3.

31. Röösli E, Rice B, Hernandez-Boussard T. Bias at warp speed: how AI may contribute to the disparities gap in the time of COVID-19. *J Am Med Inform Assoc* 2020; ocaa210. Online ahead of print.

32. Hooper MW, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA* 2020; 323 (24): 2466.

33. Wright A, Aaron S, Sittig DF. Testing electronic health records in the "production" environment: an essential step in the journey to a safe and effective health care system. *J Am Med Inform Assoc* 2017; 24 (1): 188–92.

34. ISO/IEC 15026-2:2011. Systems and software engineering—Systems and software assurance—Part 2: Assurance case. November 2011. https://www.iso.org/standard/52926.html Accessed June 11, 2020.

35. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366 (6464): 447–53.

36. Robbins R, Brodwin E. An invisible hand: Patients aren't being told about the AI systems advising their care. STAT. 15 July 15 2020. https://www.statnews.com/2020/07/15/artificial-intelligence-patient-consent-hospitals/ Accessed August 29, 2020.

37. 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 Infom Assoc* 2020; 27 (12): 2024–7.

38. Food & Drug Administration. 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 Accessed June 11, 2020.

39. American Medical Informatics Association. Comments to the FDA RE: Docket No. FDA-2019-N-1185; "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." June 2019. https://www.amia.org/sites/default/files/AMIA-Response-to-FDA-AIML-SaMD-Modifications-Draft-Framework_0.pdf Accessed June 11, 2020.

40. 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.

41. Perlin JB, Baker DB, Brailer DJ, et al. *Information Technology Interoperability and Use for Better Care and Evidence: A Vital Direction for Health and Health Care.* Washington, DC: National Academy of Medicine; 2016. https://nam.edu/information-technology-interoperability-and-use-for-better-care-and-evidence-a-vital-direction-for-health-and-health-care/ Accessed June 11, 2020.

42. Wright A, Sittig DF, Ash JA, et al. Development and evaluation of a comprehensive clinical decision support taxonomy: comparison of front-end tools in commercial and internally developed electronic health record systems. *J Am Med Inform Assoc* 2011; 18 (3): 232–42.

43. Lee JM, Hirschfeld E, Wedding J. A Patient-Designed Do-It-Yourself Mobile Technology System for Diabetes: Promise and Challenges for a New Era in Medicine. *JAMA* 2016; 315 (14): 1447–8.

44. Kesavadev J, Srinivasan S, Saboo B, et al. The do-it-yourself artificial pancreas: a comprehensive review. *Diabetes Ther* 2020; 11 (6): 1217–35.

45. Food and Drug Administration. Developing a Software Precertification Program: A Working Model v1.0. Real World Performance Analytics Framework. 2019. https://www.fda.gov/media/119722/download Accessed July 28, 2020.

46. Miller RA, Schaffner KF, Meisel A. Ethical and legal issues related to the use of computer programs in clinical medicine. *Ann Intern Med* 1985; 102 (4): 529–37.

47. Goodman KW. Ethics in health informatics. *Yearb Med Inform* 2020; 29 (01): 026– 31.

48. Association for Computing Machinery. Statement on Algorithmic Transparency and Accountability. 2017. https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf Accessed August 6, 2020.

49. 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.

50. Cosgriff C, Stone D, Weissman G, et al. The clinical artificial intelligence department: a prerequisite for success. *BMJ Health Care Inform* 2020; 27 (1): e100183.

51. CAHIIM. Health Informatics Accreditation. 2020. https://www.cahiim.org/accreditation/health-informatics Accessed July 28, 2020.