Journal of the American Medical Informatics Association, 2024, **31(11)**, 2730–2739
https://doi.org/10.1093/jamia/ocae209
Advance access publication 26 September 2024
Case Report

Case Report

Toward a responsible future: recommendations for AI-enabled clinical decision support

Steven Labkoff, MD[1,2], Bilikis Oladimeji, MBBS, MMCi[3], Joseph Kannry, MD[4], Anthony Solomonides, PhD[5], Russell Leftwich, MD[6], Eileen Koski, MPhil[7], Amanda L. Joseph, MSc[8], Monica Lopez-Gonzalez, PhD[9], Lee A. Fleisher, MD[10], Kimberly Nolen, BS, PharmD[11], Sayon Dutta, MD, MPH[12,13,14], Deborah R. Levy, MD, MPH, MS[15,16], Amy Price, DPhil[17,18], Paul J. Barr, PhD[17], Jonathan D. Hron, MD[19,20], Baihan Lin, PhD[21,22], Gyana Srivastava, BA[2,23], Nuria Pastor, MSc[24], Unai Sanchez Luque, MSc[24], Tien Thi Thuy Bui, PharmD[25], Reva Singh, JD[26], Tayler Williams, BA[26], Mark G. Weiner, MD[27], Tristan Naumann, PhD[28], Dean F. Sittig, PhD[29], Gretchen Purcell Jackson, MD, PhD[30,31] and Yuri Quintana, PhD*[,2,8,14,32]

[1]Quantori, Boston, MA 02142, United States, [2]Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA 02215, United States, [3]UnitedHealth Group, Minnetonka, MN, United States, [4]Division of General Internal Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, United States, [5]Research Institute, Endeavor Health, Evanston, IL 60035, United States, [6]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, United States, [7]IBM Research, Yorktown Heights, NY, United States, [8]School of Health Information Science, University of Victoria, Victoria, BC, Canada, [9]Cognitive Insights for Artificial Intelligence, Baltimore, MD, United States, [10]Anesthesiology and Critical Care, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, United States, [11]Pfizer Inc, New York, NY, United States, [12]Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, United States, [13]Clinical Informatics, Mass General Brigham Digital, Boston, MA, United States, [14]Harvard Medical School, Boston, MA, United States, [15]Department of Medicine, Pain Research Informatics Multimorbidities and Epidemiology (PRIME) Center, VA-Connecticut Healthcare System, West Haven, CT, United States, [16]Department of Biomedical Informatics and Data Sciences, Yale School of Medicine, New Haven, CT, United States, [17]The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, NH, United States, [18]BMJ, London, United Kingdom, [19]Department of Pediatrics, Division of General Pediatrics, Boston Children's Hospital, Boston, MA 02115, United States, [20]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, United States, [21]Departments of AI, Psychiatry, and Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, United States, [22]Berkman Klein Center for Internet and Society, Harvard Law School, Cambridge, MA, United States, [23]Harvard School of Public Health, Boston, MA 02115, United States, [24]Vitalera, Barcelona 08028, Spain, [25]Massachusetts College of Pharmacy and Health Sciences, Boston, MA, United States, [26]American Medical Informatics Association, Washington, DC, United States, [27]Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, United States, [28]Microsoft Research, Redmond, WA, United States, [29]Department of Clinical and Health Informatics, University of Texas Health Science Center, Houston, TX, United States, [30]Intuitive Surgical, Nashville, TN, United States, [31]Department of Pediatrics and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States, [32]Homewood Research Institute, Guelph, ON, Canada

*Corresponding author: Yuri Quintana, PhD, Division of Clinical Informatics, Beth Israel Deaconess Medical Center, 133 Brookline Avenue, HVMA Annex, Suite 2200, Boston, MA 02215, United States (yquintan@bidmc.harvard.edu)

Abstract

Background: Integrating artificial intelligence (AI) in healthcare settings has the potential to benefit clinical decision-making. Addressing challenges such as ensuring trustworthiness, mitigating bias, and maintaining safety is paramount. The lack of established methodologies for pre- and post-deployment evaluation of AI tools regarding crucial attributes such as transparency, performance monitoring, and adverse event reporting makes this situation challenging.

Objectives: This paper aims to make practical suggestions for creating methods, rules, and guidelines to ensure that the development, testing, supervision, and use of AI in clinical decision support (CDS) systems are done well and safely for patients.

Materials and Methods: In May 2023, the Division of Clinical Informatics at Beth Israel Deaconess Medical Center and the American Medical Informatics Association co-sponsored a working group on AI in healthcare. In August 2023, there were 4 webinars on AI topics and a 2-day workshop in September 2023 for consensus-building. The event included over 200 industry stakeholders, including clinicians, software developers, academics, ethicists, attorneys, government policy experts, scientists, and patients. The goal was to identify challenges associated with the trusted use of AI-enabled CDS in medical practice. Key issues were identified, and solutions were proposed through qualitative analysis and a 4-month iterative consensus process.

Results: Our work culminated in several key recommendations: (1) building safe and trustworthy systems; (2) developing validation, verification, and certification processes for AI-CDS systems; (3) providing a means of safety monitoring and reporting at the national level; and (4) ensuring that appropriate documentation and end-user training are provided.

Received: February 25, 2024; **Revised:** July 8, 2024; **Editorial Decision:** July 22, 2024; **Accepted:** August 13, 2024

Discussion: AI-enabled Clinical Decision Support (AI-CDS) systems promise to revolutionize healthcare decision-making, necessitating a comprehensive framework for their development, implementation, and regulation that emphasizes trustworthiness, transparency, and safety. This framework encompasses various aspects including model training, explainability, validation, certification, monitoring, and continuous evaluation, while also addressing challenges such as data privacy, fairness, and the need for regulatory oversight to ensure responsible integration of AI into clinical workflow.

Conclusions: Achieving responsible AI-CDS systems requires a collective effort from many healthcare stakeholders. This involves implementing robust safety, monitoring, and transparency measures while fostering innovation. Future steps include testing and piloting proposed trust mechanisms, such as safety reporting protocols, and establishing best practice guidelines.

Key words: clinical decision support; artificial intelligence; clinician AI competencies; patient safety; algorithmic transparency.

Introduction

In the early stages of clinical informatics, clinical decision support (CDS) systems were some of the earliest applications of computers in healthcare.[1,2] The integration of computing technology into medical practice in the 1960s and 1970s facilitated the advent of computer-assisted decision-making in areas ranging from antibiotic selection to the management of acid-base imbalances.[3] The subsequent decline in memory costs and the doubling every 2 years of computational capacity led to the expansion of hospital-wide CDS systems through the 1980s and 1990s. This period was characterized by the emergence of more sophisticated expert systems, such as the Quick Medical Reference,[4] DXPlain,[5] and ILIAD,[6] designed to augment clinical diagnosis. Efforts like Rind's renal failure studies sought to institute real-time clinical alerts.[7] Refining algorithms and integrating detailed patient data, including prior serum creatinine levels, age, and concurrent medications, improve CDS recommendations' precision and contextual relevance.

Early CDS could provide advice on making a diagnosis or picking a therapy. They accomplished this through algorithms, which could be comprehended using logistic regression, recursive partitioning, and other mathematical and computational methods. Yet the underlying data used to train these systems was rarely exposed for review. It was difficult, if not impossible, to tell if a system was trained on a dataset with intrinsic bias. However, the ability to cause harm by mistake in use or output due to bias was largely constrained by the limited availability of these systems to a relatively small set of individuals who knew how to use them and the patients they treated.

By the early 1990s, it was recognized that further transparency was needed. Miller and Gardner[8] provided an extremely comprehensive review of the domain that outlined the various levels of complexity and risk associated with decision support systems of the day. They provided a comprehensive approach to document a variety of parameters of complexity, which, in turn, reflected the degree of review and oversight that would be needed to ensure the safe deployment and use of such systems. The evolution of systems and the democratization of data led to the growth of CDS deployment.

By 2017, Labkoff and Sittig identified challenges to maintaining the safety and quality of data used in CDS in clinical care.[9] They argued that there should be ongoing oversight and a way for patients who have suffered harm due to the use of CDS to report these issues to a clearinghouse at the federal level. With the dramatic growth in the capabilities and diversity of AI-CDS, the need for such a reporting mechanism to promote the safety, transparency, and trust of these systems is now even more urgent.

Today, we are witnessing an ongoing expansion in machine learning and artificial intelligence (AI) capabilities. Notably, with the introduction of large language models in November of 2022, interest in integrating AI into CDS was rekindled. The 2023 HIMSS conference showcased the proliferation of AI in healthcare systems, an exhibition unparalleled in previous years.

However, the historical challenges of algorithm development, validation, and deployment, including the representativeness of the training data, have been amplified in the current landscape. Today, despite privacy concerns, the creation and utilization of extensive datasets for training AI-CDS systems are commonplace. Similar to their predecessors, these systems may harbor latent biases that are extremely difficult to account for. The ease of acquiring vast training datasets in the modern era exacerbates the challenges previously encountered, necessitating heightened scrutiny of AI-CDS systems' training data and deployment mechanisms. Additionally, there is even less transparency in how decisions, advice, and recommendations are made. Because there is often no way to decipher the pathway to a given recommendation, the systems that produce them are often called "black boxes."

While established methodologies, such as randomized controlled trials (RCTs), have been used to evaluate the effectiveness and safety of traditional CDS systems,[10] these approaches may not be sufficient for AI-enabled CDS (AI-CDS). The dynamic nature of AI algorithms, which can evolve and adapt over time through continuous learning and model updates, introduces the challenge of model drift. As a result, new evaluation approaches are needed to assess the performance, safety, and robustness of AI-CDS systems throughout their lifecycle, accounting for potential changes in model behavior.

Generative AI technologies and extensive language models (LLMs) like ChatGPT represent a significant advancement over traditional machine learning approaches. Traditional machine learning models typically rely on structured datasets and are designed to perform specific tasks, such as classification or regression, based on predefined features. These models require labeled data and often involve feature engineering to optimize performance. The critical differences between LLMs and traditional machine learning approaches include LLMs are trained on extensive, diverse datasets from various sources, while traditional models use smaller, domain-specific datasets. LLMs utilize unsupervised learning to understand language patterns and generate text, whereas traditional models rely on supervised learning with labeled data. LLMs can generate new text based on input prompts, making them versatile, while traditional models are designed for specific predictive tasks. LLMs use large, undisclosed datasets, complicating bias assessment, whereas traditional models allow for more straightforward bias evaluation and mitigation.

Regulatory and professional bodies, such as the Food and Drug Administration (FDA),[11] American Medical Association (AMA),[12] World Health Organization (WHO),[13] European

Commission in collaboration with the European Medicines Agency (EMA),[14] and AMIA,[15] have established initial frameworks to evaluate and monitor AI systems for "fitness for purpose," efficacy, safety, and fairness. The term "fitness for purpose" refers to the ability of an AI system to effectively and efficiently fulfill the specific task or purpose for which it was designed. This includes the system's accuracy and speed and its adherence to ethical standards such as fairness, transparency, and respect for privacy. Prior literature[16] examined algorithm auditing methodologies and protocols for tailoring traditional clinical evaluation approaches to AI. However, difficulties persist in comprehensively assessing AI-CDS tools outside limited research contexts.[17,18] Moreover, healthcare stakeholders urgently need coordinated action to implement layered AI governance solutions to address issues across the entire AI-CDS lifecycle,[19] including relevant ethical and social issues, not least privacy, transparency, accountability, and equity.

Recent publications have emphasized the importance of addressing algorithmic bias,[20,21] fairness[22] issues, and the need for explainable AI and interpretability[23] in healthcare applications. Moreover, researchers have focused on developing and applying evaluation and validation methodologies to assess the performance and safety of AI-enabled CDS systems. Studies have also explored the organizational, technical, and human factors[24] that influence the adoption of AI-enabled CDS. There is a need to achieve consensus for the evaluation of AI for CDS. This paper will provide the findings and recommendations of a detailed consensus process with a broad range of stakeholders to ensure that AI-CDS is safer, more effective, and more equitable in a way that represents the multiple viewpoints of the key stakeholder groups.

Materials and methods

During the summer of 2023, in partnership with the American Medical Informatics Association (AMIA), the DCI Network,[25] a consortium created by the Division of Clinical Informatics (DCI) at Beth Israel Deaconess Medical Center (BIDMC), convened and created a working group focused on AI in healthcare. We prepared 4 antecedent webinars on "AI in Healthcare" that covered topics including (a) Bioethics and Religion, (b) Patient perspectives, (c) Real-world Data and the Regulatory Perspective, and (d) Risk management, Trust, and Liability.[26]

A month later, we convened a cross-functional, interdisciplinary team involving over 200 experts from government, standards organizations, private industry, ethicists, patients and advocates, religious leaders, clinical informaticians, and regulators, gathered through 4 webinars and a 2-day live consortium at the Harvard Medical School in September 2023.[27] Over these 2 days, the group heard from multiple speakers and, later, broke out into working groups. We focused on 3 specific domains of how AI is being deployed in the healthcare ecosystem: (1) a patient's viewpoint on how they are using AI in their healthcare journeys; (2) how AI is impacting CDS; and (3) how AI is being used to create, enrich, and use real-world data/evidence. The participants were tasked with identifying challenges, proposing strategies, and creating key goals to ensure that AI-CDS can be built, used, and evaluated safely, efficiently, and equitably in the context of both the work that has been done in the field, as well as how to address current-day and future challenges presented by the democratization of AI-CDS systems. The sub-groups developed rapid brainstorming methods and group discussions to develop the core ideas and challenges. Over the next 4 months, a sub-group of the team (the authors) met in weekly online discussion meetings to develop consensus recommendations presented in this paper. The group used a Delphi approach to iterate on the core challenges and the consensus approaches for how they should be prioritized and addressed. This paper summarizes the findings and recommendations of this process and the implications for the healthcare field.

Results

Governance of these interconnected issues requires multipronged initiatives around transparency, training, and infrastructure from developers, regulators, and healthcare delivery organizations. The key findings are shown in Table 1 and can be grouped into 4 primary domains: (1) To build safe and trustworthy systems by creating a systematic approach to a "nutrition label"; (2) Validation and verification (ie, certification of AI-CDS systems)—create a systematic approach to testing and validation; (3) Safety monitoring and reporting—the creation of an AI-CDS reporting center for adverse events; and (4) Documentation and end-user training—provide comprehensive and certified training for professional and lay users of AI-CDS.

Recommendations

Building safe and trustworthy systems

Each domain in Table 1 can be reflected in the various stages of the AI-CDS development lifecycle. Figure 1 presents a virtuous cycle, leveraging design thinking concepts in software development to visualize the stated challenges.[28] Several frameworks have been proposed for AI.[29,30] This illustration outlines a typical development and deployment lifecycle. By using trusted standards for system development, we can ensure that best practices are adhered to, that systems can support interoperability, and facilitate workflow integration.

Integration into existing systems: the role of international consensus standards

One way to drive trust in AI-CDS in clinical care is to provide a framework and standards to promote transparency and interoperability. To facilitate the consistent adoption of such tools, healthcare standards development organizations (ie, HL7, IEEE, EN TC 251 - Technical Committee on Health Informatics Europe, and International Organization for Standardization [ISO] Technical Committee 215) should strive to define specifications around required fields, formats, and APIs for data and metadata access and use. The unique AI Identifier (UAII) label (see below) could be adapted as a new class within the HL7 FHIR standard, enabling seamless integration with other clinical and research data. Other collaborative efforts tapping into diverse expertise have successfully propagated terminology system standards like SNOMED CT through SNOMED across care settings.[31,32]

However, establishing standards alone is insufficient to gain the necessary trust in AI-CDS, especially as these systems may eventually undertake tasks such as prescribing medications without human oversight. Although this scenario is not yet a reality, it represents a conceivable future direction for AI-CDS. Incorporating these unattended AI-CDS

Table 1. Key challenges identified by the consensus group (listed in alphabetical order).

Core concept	Challenge
Adverse events	Lack of real-time monitoring and reporting to enable aggregation, analysis, and alerts around potential safety issues.
Bias	AI algorithms and subsequent recommendations may reflect both dataset and development biases due to the following: 1) Differential representation of race, gender, socioeconomics, clinical subpopulations, and health disparities in datasets. 2) Inaccurate design assumptions due to experiential or other biases by designers, developers, or users. 3) Failure to consider the values and preferences of the patient. 4) Differential coding of data due to different local conventions or interpretations of coding standards leads to inaccurate or inappropriate data aggregation and analysis.
Credibility and explainability	Current AI systems provide limited explanations or citations for their reasoning.
Documentation and transparency	Insufficient clarity regarding the provenance of biases in training data, model development, and real-world testing procedures may undermine the credibility and adoption of AI tools.
Generalizability and portability	The specificity of the clinical context and the data used for training may limit generalizability. Application to new or evolving scenarios can lead to unpredictable results in other settings.
Inaccurate inferences	Inadequately constrained AI models may provide false or unsupported recommendations. Inadequately constrained AI models refer to those AI systems that lack sufficient limitations or rules in their design and operation. These models may not have the necessary restrictions to guide their learning process effectively. This leads to overfitting, where the model learns the training data too well and performs poorly on new, unseen data.
Integration	Interoperability challenges may hamper AI integration with data sources and clinician workflows; they can affect the ability to get data in and make recommendations actionable.
User competency	"Users" primarily refers to clinicians interacting with the AI-enabled Clinical Decision Support (AI-CDS) systems. Ensuring clinicians have the necessary competencies to use these systems effectively is crucial for patient safety and the system's overall efficacy. Inadequate training or demonstrated competency in assessing recommendations and deploying AI-CDS risks patient harm and distrust.
Validation and certification	Frequent model retraining on new data, which is required to prevent model "drift," creates moving targets for validation, benchmarking, and auditing. The existing regulatory apparatus is not equipped for the technical and governance challenges of near real-time review, re-evaluation, and certification.
Veracity of results	Variability in recommendations or detection of AI errors from an AI-CDS will erode confidence in the system. Establish clear criteria for veracity based on rigorous testing, validation, and continuous performance monitoring.

interventions within existing EHRs without sufficient testing and trust will open new, complex trust challenges.

Clinical integration and decision partnerships

For sustainable, safe, and effective adoption of AI-CDS in healthcare settings, it must be part of the clinical workflow and integrated seamlessly into HIT, such as EHRs, clinical information systems, personal health records, and patient portals. AI-CDS should also help diminish documentation challenges prevalent in EHR systems today.[33,34]

Such systems should be designed to provide user-friendly summaries for both clinicians and patients and support documentation encompassing:

- The goal and specific questions asked of the AI.
- The AI-generated response.
- Whether or not a clinician reviewed and agreed with the recommendation.
- The date and time when the system was used.
- Any references needed to clarify how and when the system was used.

Equity

The WHO defines equity as the absence of unfair, avoidable, or remediable differences among groups of people, whether those groups are defined socially, economically, demographically, or geographically, or by other dimensions of inequality (eg, sex, gender, ethnicity, disability, or sexual orientation).[35] Moreover, such disparities often originate in historical and systemic care delivery contexts, clinical research, and evidence generation.

As such, in developing and adopting AI-CDS in healthcare, it is essential to contextualize it as a facilitator as well as a potential barrier to equitable healthcare delivery.[36–38] To avoid creating and perpetuating new dimensions of disparities, individual researchers, agencies, professional bodies, and the government have proposed various frameworks for ensuring health equity is considered in all innovations, including those driven by AI/ML.[39,40]

An article by the Research Triangle Institute summarizes the 8 primary health equity frameworks and measures that inform how health equity performs across the Institute for Health Improvement (IHI), the Robert Wood Johnson Foundation (RWJF), the Centers for Medicare and Medicaid Services (CMS), the National Committee for Quality Assurance (NCQA), and the Joint Commission International (JCI). Concepts such as 'Inclusion by design'[41] and 'algorithm vigilance',[42] among others, enhance fairness in system development, deployment, and evaluation. As part of the concept of Inclusion by Design, the Global Future Council on Artificial Intelligence for Humanity designed an Inclusive AI Lifecycle.[34] This provides a systems view of embedding the goal of equity in all steps of the AI development lifecycle and mapping both the builder and stakeholder/governance ecosystem.

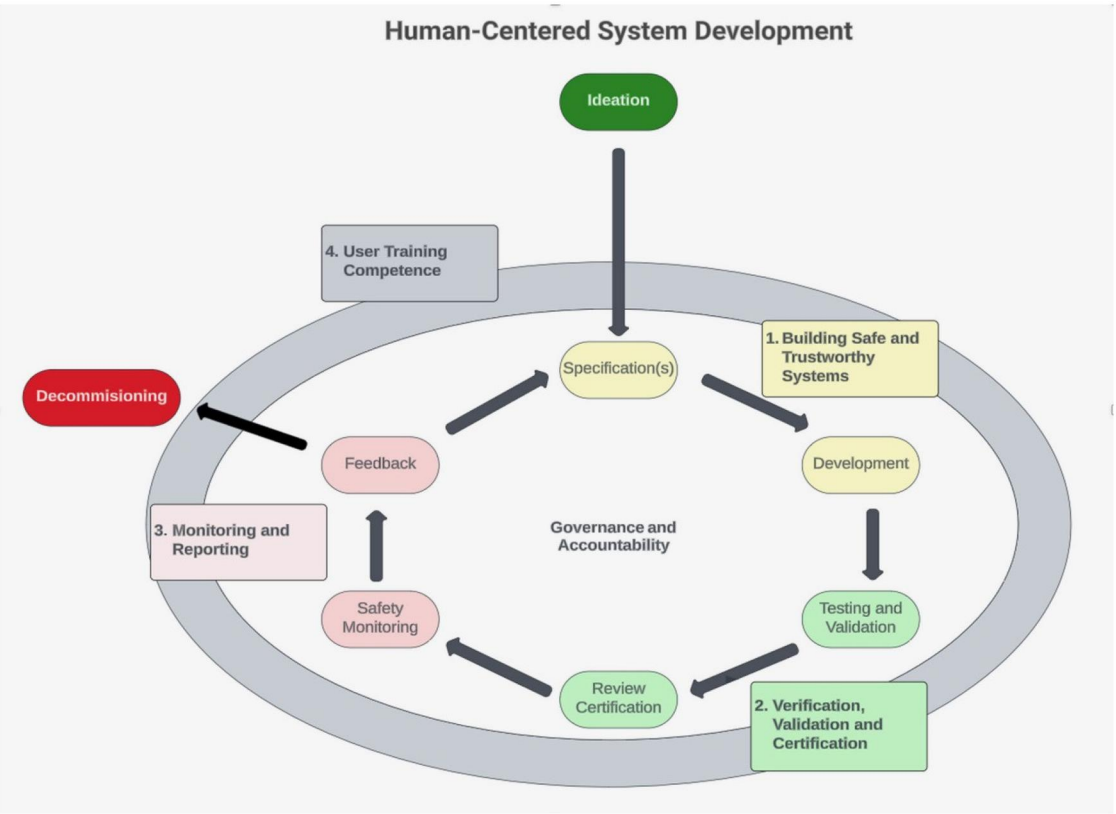

Figure 1. Human-centered AI development cycle.

Themes across these equity frameworks include the following:

- Governance and organizational support
  There are business, moral, and health imperatives for making equity a strategic and foundational goal of all projects or programs in healthcare, which requires strong organizational support and governance.[43]
- Data **management**
  AI-CDS systems depend on the datasets used to train the algorithms and make recommendations. Furthermore, several data characteristics should be assessed and documented to guide model developers and users, including provenance, representation, context, and assumptions.[44]
- **Development and Implementation Team Requirements**
  - **Assessment of disparities and social determinants of health:** One must understand how sociodemographic determinants of health may affect outcomes and implement appropriate detection and mitigation strategies.
  - **Inclusive stakeholder/community engagement:** Must be multidisciplinary (ie, clinical, technical, operational, and end-users). All stakeholders should be involved in problem definition, validation, and deployment planning throughout the AI development cycle (Figure 1).
  - **Diverse workforce:** The workforce should be diverse across multiple dimensions of identity, ideally representing the population on which the model is intended to be deployed as clinicians or patient end-users. Diversification in teams could provide varied perspectives on population samples that are considered representative diverse views on ethical principles, enhancing the accuracy of model development and mitigating the risks of introducing implicit bias into such systems.
- **Audit and evaluation of health outcomes**
  It is crucial to monitor and assess models, recommendations, and outcomes considering various forms of identity and social factors, such as race, socioeconomic status, health literacy, insurance coverage, etc. These factors should improve over time as the technology is developed and used (Figure 1).

We recommend incorporating best practices and principles from health equity frameworks[45] and research through the entire AI-CDS (Figure 1) system lifecycle to reduce disparities and promote equitable care.[46]

Verification, validation, and certification

"Verification" refers to ensuring the accuracy, reliability, and safety of AI algorithms designed for medical applications, including rigorous testing and assessment of the AI's performance against predefined benchmarks. Data validation assesses the quality and completeness of the data used to train and test the AI model; algorithmic validation assesses the model's ability to produce consistent and accurate results; and clinical validation evaluates the efficacy and reliability of the results in a clinical context.

Certification refers to having the AI-CDS' output (recommendations) evaluated against some agreed-upon standard from either a standards development organization or regulatory body. Certification helps to ensure that the system meets

Table 2. Verification, validation, and certification issues for verification of AI-CDS.

Model training	• Documentation of data origins and characteristics is critical to the credibility of advice or direction provided by these systems
	• Identification and documentation of known biases in the training datasets:
	  • Type of data (eg, EHR data, claims data)
	  • Its origin, where it was collected (facility, geography), and the nature of the population from which it was gathered (demographics)
	  • Any known intrinsic biases in the training data (eg, limited population demographics)
	• Identification of training methods (eg, deep learning, supervised learning, unsupervised learning, etc)

Explainability[47]	• How have the decisions and recommendations been reached?
Use-case dependency	• Virtually all AI-CDS systems are optimized for a particular use case and population. Results may be unpredictable if applied to different use cases or populations.
Veracity of results	• Variability in recommendations or detection of AI errors from an AI-CDS will erode confidence in the system.
End-user training	• The ability to interpret results will depend on the operator's knowledge, experience, and training. Proper training should be clearly defined and repeated annually or in response to system updates and changes.

some known standard of accuracy, which helps foster trust. Table 2 identifies additional verification, validation, and certification issues that must be addressed to verify that AI-CDS provides trusted results.

AI-CDS documentation

One of the first recommendations from our working group was creating a transparency label for AI-CDS systems called the Unique AI Identifier (UAII). Its purpose is to address trust and enhance transparency by disclosing relevant aspects of such systems. The UAII is similar to a nutrition, implantable device, or medical product label. It provides vital metadata needed to use AI-CDS systems safely, responsibly, and with accountability. It also provides a complimentary means of maintaining a certification cycle for a given system under review.

We recommend a set expiration of certification when a UAII must be reviewed, along with any data about the system's use, performance, adverse events, or other outcomes.

Sendak et al[48] suggested one example of such a label in 2020. This was modeled after the medical label. In addition to Sendak's recommendations, our consensus working group had additional fields to be considered for the UAII, as found in Table 3.

Monitoring and reporting

To ensure safe use of AI-CDS, there is a need for the creation of a National Healthcare AI-CDS Safety Reporting Clearinghouse. The FDA maintains the MedWatch system as a clearinghouse for adverse event reporting for biologics, prescription and over-the-counter drugs, combination products, medical devices, special nutrition products, cosmetics, and food.[50] This system enables medical professionals and others to report adverse events promptly and comprehensively. The system also provides real-time safety information alerts via a subscription email list. The FDA is also involved with AI in post-market surveillance of Software as a Medical Device (SaMD).[51] However, some important questions remain.[52] The FDA also maintains the MAUDE[53] for tracking issues, specifically with medical devices. Having something similar for AI-CDS event reporting would be the next logical step.

An AI-CDS clearinghouse must facilitate and standardize reporting of unexpected or adverse events that stem from using AI-enabled CDS and must become part of the public record. Adequate funding should be provided so the agency can monitor and report the reported information.

User competence (training)

End-users require training as with any new medical guideline, device, procedure, or medical therapy. As AI-CDS becomes mainstream, system- and tool-specific training of end users is critical for overall safe use, regardless of whether or not it resides inside otherwise familiar EHR systems.

Health professions educators should define new competencies for using AI-CDS in clinical practice, and these training requirements should be incorporated into medical education programs at all levels and evolve with the technology.[54] We propose that, as part of the UAII process, developers, manufacturers, and content authors specify appropriate training requirements for users. This should include competencies such as the context of the model, assumptions made, and the model's limitations. The actual training should be up to experts in adult education to provide such training and certification.

Our working group recommends creating a similar set of constructs and a parallel system for training requirements, which should be mandatory (depending on the application's envisioned end users—for clinicians or laypeople) and documented and monitored as part of overall risk mitigation and legal/compliance constructs.

Discussion

AI-CDS promises to transform healthcare decision-making, necessitating investments in transparency, monitoring, and oversight. This consensus analysis establishes a framework emphasizing trustworthy systems, verification, validation, certification, monitoring, reporting, documentation, and training. Key concepts include Unique AI Identifier labels (UAII), model training, explainable AI, result veracity, and adverse event reporting.

In recognizing that some failures of CDS systems are inevitable, it is essential to design them to be as fail-safe as possible. This involves comprehensive testing and validation, user-centered design, and a layered decision-making process to ensure reliability and safety. Predictable errors such as data entry inaccuracies, algorithm bias, clinical context mismatches, and model drift can be mitigated through strategies such as automated data entry, continuous bias assessment, incorporating contextual data, and regular model updates.

Table 3. Key fields to be included in the UAII, in descending order of importance to the clinician.

Optimal use case	Descriptions of circumstances and situations where the model can be best used.
Information about training data and potential biases	A comprehensive review of the data used to train the system, encompassing information on how and where it was obtained, population, location, care setting (if appropriate), and date range. Any known imbalances or biases are to be listed. Measures taken to clean the data: how missingness was addressed, how representativeness was evaluated, how the data were de-biased or the effects of bias controlled for, and how new data types were incorporated.
Duration of certification	Systems should be certified for use for a finite time to ensure they continue providing valid, accurate, and safe recommendations. All systems should be revalidated after system or training data updates. Adaptive AI systems and those that self-modify must be reviewed and calibrated regularly to correct any drift.
Version information	Clear identification of the current version number or other means of version identification must be provided and refreshed during updates. Version disclosure should not be limited to the software but to the data and the training data.
Conflict of interest disclosures	Standard processes for conflict-of-interest disclosures should apply to all system developers.
Limitations and warnings	The UAII should enumerate any known limitations and warnings regarding domains of use, use cases, and any other information that could influence overall use or interpretations.
Algorithms employed	Disclosures about the types and sources of algorithms employed by the AI-CDS.
Validation and certification	The UAII should show evidence of how the system was tested and how the results of those tests were validated, including unexpected outcomes.
Recommended training	Any prerequisites needed to certify the user should be disclosed in the UAII. For instance, if a licensed cardiologist should be the only person using an AI-CDS system, UAII should make this clear. Conversely, if it can be used by untrained end-users (ie, patients, laypersons), that too should be clearly described.[49]

By implementing these guidelines, we can enhance the integration of CDS into clinical care, thereby improving patient safety and the overall efficacy of clinical decision-making.

These new systems can be modified in many ways, ranging from updates to training algorithms to updating their training datasets. A system to test the continued validity and credibility of recommendations after these changes is essential. Even small changes in these systems can potentially affect how they arrive at recommendations. The current regulatory tools at our disposal are not sufficiently agile to meet these challenges.

Evaluation of AI-CDS efficacy and safety traditionally relies on randomized controlled trials (RCTs), cohort, and case-control studies. However, the dynamic nature of nondeterministic algorithms requires new approaches. Unique to AI-CDS, there must also be capped autonomy levels (eg, at what point, if ever, does the software become autonomous, not requiring human intervention and confirmation).

Phased-in changes to medical education curricula and competency testing for AI use in clinical environments will be needed. Patients' understanding and reaction to these systems also require careful consideration. Operationalizing these recommendations involves overcoming socio-technical barriers to explainability and encouraging transparent developer behaviors while avoiding overly restrictive policies.

AI systems' modifiability necessitates continuous validity testing. Current regulatory tools may lack the agility to address these challenges effectively. Documenting intrinsic biases in AI models is crucial for informed utilization. Underlying disparities existing today in the real world have the potential to be promulgated in AI training models.[55] Correcting embedded, intrinsic bias in existing real-world data may be impossible. However, documenting such issues (in the UAII) is essential to informing appropriate utilization—or when not to use such models.[37]

Public policy support is crucial for the proposed framework's consistent implementation. This may involve executive orders, federal regulations, or legislation. The FDA has developed a framework for Software as Medical Device (SaMD) and has considered regulating AI-CDS output since 2019 and has already developed a framework for using SaMD as CDS[56] but challenges persist regarding AI decision-making explainability.[57]

AI systems currently do not consistently provide a plausible explanation as to how or why recommendations are given (ie, "black box" system). As demonstrated with Meaningful Use, Stage 2 in 2012,[58] there is a need to provide the rationale to its recommendations.[59] While voluntary UAII creation might lead to uneven adoption, mandating UAII publication could promote widespread use. However, a balance is needed between standardization and avoiding innovation-stifling mandates.

The slow process of the development of national laws and policies requires more pragmatic approaches that could be implemented sooner. Industry self-regulation through Standards Development Organizations such as HL7, IEEE, and ISO[60] could drive consistent adoption of transparency tools such as UAII. Standardizing specifications for data fields, formats, and APIs could enable seamless integration with clinical and research data. While collaborative efforts have successfully propagated standards like SNOMED CT, challenges remain, as seen with varying implementations of HL7 2.x by different vendors.[61,62]

AI-CDS tools introduce new challenges in medical informatics, particularly in ensuring trust.[63] While UAII labels are helpful, trust requires rigorous validation and verification. Standard methods may not account for issues like model drift post-launch. A multi-layered approach should combine internal testing under varied use cases and ongoing monitoring (eg, using the UAII above label to complement adverse event reporting).[52] Conversely, there is a risk of over-reliance on AI-CDS recommendations, as observed in some studies.[64]

Legal and liability challenges are currently being debated, especially with respect to autonomous AI CDS.[65] CDS was faced with similar questions and addressed them best by clearly stating that there was a human in the loop.[66] What happens if clinicians trust AI too much and are asked to justify decisions based on logic they don't understand? What

happens if no human is in the loop, such as with autonomous AI? Much of this remains to be examined, let alone decided.

The cliché "data is the new oil" highlights data's importance in powering AI, akin to oil's role in traditional industries. However, it also underscores the privacy risks associated with data used in AI systems. As these systems evolve, the need for robust data privacy measures intensifies. Thus, the phrase serves as a reminder of data's dual role in AI: a valuable innovation driver and a potential privacy risk. New legislation like the European AI Act[67] and the USA's Whitehouse Blueprint,[68] for an AI Bill of Rights may help to improve privacy and safety. However, frameworks like the one proposed in this paper are needed to guide the development and evaluation of AI-CDS systems so they comply with the various emerging regulations.

Fairness in algorithmic CDS systems is complex, with various metrics often leading to conflicting results. The choice of fairness metric should align with the specific clinical purpose. For example, equalized odds may be prioritized in diagnostics, while predictive parity might be essential for treatment recommendations. To navigate these complexities, we recommend context-specific evaluation, stakeholder engagement, hybrid approaches balancing multiple metrics, and maintaining transparency and flexibility. These principles can ensure fairness in CDS systems is tailored to clinical applications, enhancing equity and efficacy in patient care.

Data privacy and the balance between data dissemination and model quality are crucial in AI-enabled CDS systems. While disseminating healthcare data can enhance AI model performance, it raises patient privacy and data security concerns. Federated learning offers a potential solution by training AI models on decentralized data. However, implementing federated learning in healthcare presents challenges, including standardized data format requirements, potential model bias, and computational resource needs. Researchers, healthcare institutions, and policymakers must carefully balance data privacy and AI model quality, constructing robust frameworks for responsible data sharing and federated learning application in healthcare.

Records scattered across multiple EHR systems and in various formats complicate AI-CDS integration. Proposed solutions include adopting healthcare data standards like HL7 FHIR; developing advanced data integration strategies; investing in data mapping to convert unstructured data; implementing pilot programs; and fostering collaboration among stakeholders to create a supportive ecosystem for AI-CDS integration.

Re-certification and evaluation of AI systems for CDS are crucial for ongoing efficacy and safety. AI models may experience "drift" over time due to changes in data distribution, training data updates, self-modifications, or patient population shifts. Regular review, calibration, and revalidation are necessary to correct drift and maintain accuracy. Certifying systems for finite periods ensures periodic reassessment, safeguarding against outdated or erroneous recommendations. These processes are integral to AI's responsible deployment and maintenance in healthcare.

AI-enabled CDS systems require a rigorous approval process overseen by regulatory bodies like the FDA before effective monitoring and adverse event reporting can be implemented. This process includes early FDA dialogue, comprehensive clinical validation studies with RCTs, in-depth risk assessment and mitigation strategy development, detailed approval submission with all validation data and risk assessments, and post-market surveillance for monitoring real-world performance and identifying adverse events through continuous monitoring and reporting mechanisms.

In summary, AI tools are poised to transform medicine by integrating into CDS. As AI technologies proliferate, their full integration into clinical workflow and healthcare delivery is inevitable. Urgent standardization of frameworks, implementation of guiding principles, and establishment of best practices are needed to safeguard medical practice and patient outcomes. Given the prevalence of preventable adverse events in hospital admissions,[66] we are optimistic that AI-CDS can lead to better outcomes we are optimistic that AI-CDS can significantly improve healthcare outcomes.

Conclusion

The speed and breadth by which AI can impact CDS have become more significant due to data availability, computing power, and new LLMs; however, the risks have never been higher for unvalidated systems. In this paper, we provided specific and pragmatic approaches to improve how we evaluate AI-CDS to improve safety, efficacy, and equity. This process for reaching a consensus among many stakeholders laid out practical building blocks and infrastructure oversight to help the responsible use of AI in clinical decision-making settings. By embracing collaborative governance through principles including (1) building trustworthy and safe systems; (2) verification, validation, and certification; (3) monitoring and reporting; and (4) user competence, healthcare systems can benefit from improved efficiency, accuracy, and consistency while mitigating risks from rapidly evolving algorithms.

Acknowledgments

We thank Dr David Bray, Dr Tiffani J. Bright, and Dr Isaac Kohane for their keynote talks at the conference. We thank Dr Bray for his initial input in early discussions on the framing of this paper. We would also like to thank Rabbi Daniel Cohen and Reverend Greg Doll for their webinar on Bioethics and AI, Dr Luke Sato, Dr Paul R. DeMuro, and Kenneth E. White, JD, for their webinar on AI in Healthcare: Risk Management, Trust, and Liability, Dr Amy Price and Dave deBronkart (ePatient Dave) for their webinar on AI in Healthcare: The Patient Perspective, and Dr Anne-Marie Meyer, Mark Shapiro, and Rob Stolper for their webinar: AI In Healthcare: Real World Data Generation and The Regulatory Perspective. Lastly, we thank Dr Charles Safran for his input on the paper's later versions and his guidance at the DCI Network.

Author contributions

Yuri Quintana and Steve Labkoff (Conceptualization), Yuri Quintana and Steve Labkoff (Methodology), Yuri Quintana, Steve Labkoff, Joseph Kannry, Eileen Koski, Bilikis Oladimeji, Anthony Solomonides, and Gyana Srivastava (Writing—Original draft), All authors (Writing—Review & editing), Yuri Quintana and Steve Labkoff (Supervision), and Gyana Srivastava (Project administration)

Funding

None declared.

Conflicts of interest

S.L. is an employee of Quantori, LLC. B.O. is an employee of UnitedHealth Group. R.L. is a full-time employee of InterSystems Corporation. E.K. owns stock in IBM. L.A.F. is a Principal at Rubrum Advising, which advises companies on coverage of new technologies, including AI. K.N. is an employee and shareholder of Pfizer Inc. T.N. is an employee of Microsoft Corporation. D.F.S. is an owner of Informatics Review LLC. G.P.J. is an employee of Intuitive Surgical. G.P.J. owns stock and/or stock options for IBM, Kyndryl, and Intuitive Surgical. G.P.J. was the Past President and Past Board Chair of the American Medical Informatics Association. The authors had no other conflicts or competing interests to disclose.

Data availability

All data used in this article have been reported in this paper. There are no other data associated with this paper.

Disclaimer

The contents of this manuscript represent the view of the authors and do not necessarily reflect the position or policy of the U.S. Department of Veterans Affairs or the United States Government.

References

1. Collen MF. Computer medicine: its application today and tomorrow. *Minn Med*. 1966;49(11):1705-1707.
2. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. *Yearb Med Inform*. 2016;25(Suppl 1):S103-S116. https://doi.org/10.15265/IYS-2016-s034
3. Bleich HL. Computer-based consultation: electrolyte and acid-base disorders. *Am J Med*. 1972;53(3):285-291.
4. Parker RC, Miller RA. Creation of realistic appearing simulated patient cases using the INTERNIST-1/QMR knowledge base and interrelationship properties of manifestations. *Methods Inf Med*. 1989;28(4):346-351.
5. Barnett GO, Cimino JJ, Hupp JA, et al. DXPlain: an evolving diagnostic decision-support system. *JAMA*. 1987;258(1):67-74.
6. Lincoln MJ, Turner CW, Haug PJ, et al. ILIAD training enhances medical students' diagnostic skills. *J Med Syst*. 1991;15(1):93-110.
7. Rind DM, Safran C, Phillips RS, et al. Effect of computer-based alerts on the treatment and outcomes of hospitalized patients. *Arch Intern Med*. 1994;154(13):1511-1517.
8. Miller RA, Gardner RM. Recommendations for responsible monitoring and regulation of clinical software systems. *J Am Med Inform Assoc*. 1997;4(6):442-457.
9. Labkoff SE, Sittig DF. Who watches the watchers: working towards safety for EHR knowledge resources. *Appl Clin Inform*. 2017;8(2):680-685. https://doi.org/10.4338/ACI-2017-02-IE-0032
10. Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med*. 2012;157(1):29-43. https://doi.org/10.7326/0003-4819-157-1-201207030-00450
11. Artificial intelligence and machine learning in software as a medical device. (n.d.). Accessed January 16, 2024. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device
12. Advancing health care AI through ethics, evidence and equity. (n.d.). Accessed January 18, 2024. https://www.ama-assn.org/practice-management/digital/advancing-health-care-ai-through-ethics-evidence-and-equity
13. WHO outlines considerations for regulation of artificial intelligence for health. (n.d.). Accessed January 8, 2024. https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health
14. Regulatory framework proposal on artificial intelligence. (n.d.). Accessed January 4, 2024. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
15. Solomonides AE, Koski E, Atabaki SM, et al. Defining AMIA's artificial intelligence principles. *J Am Med Inform Assoc*. 2022;29(4):585-591. https://doi.org/10.1093/jamia/ocac006
16. Park SH, Han K, Jang HY, et al. Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis. *Radiology*. 2023;306(1):20-31. https://doi.org/10.1148/radiol.220182
17. Shah NH, Halamka JD, Saria S, et al. A nationwide network of health AI assurance laboratories. *JAMA*. 2024;331(3):245.
18. Stead WW, Aliferis C. Health AI assurance laboratories. *JAMA*. 2024;331(12):1061-1062. https://doi.org/10.1001/jama.2024.1084
19. Sittig DF, Boxwala A, Wright A, et al. A lifecycle framework illustrates eight stages necessary for realizing the benefits of patient-centered clinical decision support. *J Am Med Inform Assoc*. 2023;30(9):1583-1589. https://doi.org/10.1093/jamia/ocad122 PMID: 37414544; PMCID: PMC10436138.
20. Ferryman K, Mackintosh M, Ghassemi M. Considering biased data as informative artifacts in AI-assisted health care. *N Engl J Med*. 2023;389(9):833-838. https://doi.org/10.1056/NEJMra2214964
21. Alvarez JM, Colmenarejo AB, Elobaid A, et al. Policy advice and best practices on bias and fairness in AI. *Ethics Inf Technol*. 2024;26(2):31. https://doi.org/10.1007/s10676-024-09746-w
22. Liu S, McCoy AB, Peterson JF, et al. Leveraging explainable artificial intelligence to optimize clinical decision support. *J Am Med Inform Assoc*. 2024;31(4):968-974.
23. Vasey B, Nagendran M, Campbell B, DECIDE-AI expert group, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med*. 2022;28(5):924-933. https://doi.org/10.1038/s41591-022-01772-9. Erratum in: Nat Med. 2022;28(10):2218.
24. Wang L, Zhang Z, Wang D, et al. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. *Front Comput Sci*. 2023;5:1187299.
25. DCI Network. Boston, MA: Beth Israel Deaconess Medical Center. Accessed September 20, 2024. https://www.dcinetwork.org
26. AI in healthcare: risk management, trust, and liability. Accessed December 28, 2023. https://www.dcinetwork.org/events/190.
27. Blueprints for trust: best practices and regulatory pathways for ethical AI in healthcare. Accessed December 28, 2023. https://www.dcinetwork.org/events/192?prev_path=/events
28. Software Engineering Models and Methods Course. (n.d.). Accessed January 16, 2024. https://www.computer.org/product/education/software-engineering-models-course
29. AI guide for government: understanding and managing the AI lifecycle. (n.d.). Accessed January 18, 2024. https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/#:~:text=The%20AI%20lifecycle%20is%20the,%2C%20development%2C%20and%20deployment%20phases
30. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. Accessed January 4, 2024. https://muse.jhu.edu/article/789672/pdf
31. Shortliffe EH, Cimino JJ, Chiang MF, eds. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. 5th ed. Springer; 2021.

32. Finnell JT, Dixon BE, eds. *Clinical Informatics Study Guide: Text and Review*. 2nd ed. Springer; 2022.
33. Levy DR, Sloss EA, Chartash D, et al. Reflections on the documentation burden reduction AMIA Plenary Session through the lens of 25 × 5. *Appl Clin Inform*. 2023;14(1):11-15. https://doi.org/10.1055/a-1976-2052
34. Brender TD. Medicine in the era of artificial intelligence: Hey Chatbot, write me an H&P. *JAMA Intern Med*. 2023;183(6):507. https://doi.org/10.1001/jamainternmed.2023.1832
35. Health equity. (n.d.). Accessed January 18, 2024. https://www.who.int/health-topics/health-equity#tab=tab_1
36. Khan M, Bates D, Kovacheva V. (n.d.). The quest for equitable health care: the potential for artificial intelligence. Accessed January 10, 2024. https://catalyst.nejm.org/doi/full/10.1056/CAT.21.0293
37. Ledford H. Millions of black people affected by racial bias in health-care algorithms. *Nature*. 2019;574(7780):608-609. https://doi.org/10.1038/d41586-019-03228-6
38. Glauser W. AI in health care: improving outcomes or threatening equity? *CMAJ*. 2020;192(1):E21-E22. https://doi.org/10.1503/cmaj.1095838
39. Nordling L. A fairer way forward for AI in health care. *Nature*. 2019;573(7775):S103-S105. https://doi.org/10.1038/d41586-019-02872-2
40. Courtland R. Bias detectives: the researchers are striving to make algorithms fair. *Nature*. 2018;558(7710):357-360. https://doi.org/10.1038/d41586-018-05469-3
41. A blueprint for equity and inclusion in artificial intelligence. (n.d.). Accessed January 10, 2024. https://www.weforum.org/publications/a-blueprint-for-equity-and-inclusion-in-artificial-intelligence/
42. Embi PJ. algorithm vigilance: advancing methods to analyze and monitor artificial intelligence-driven health care for effectiveness and equity. *JAMA Netw Open*. 2021;4(4):e214622. https://doi.org/10.1001/jamanetworkopen.2021.4622
43. Donald M, Berwick MD. MPP: the moral determinants of health. Accessed January 31, 2024. https://jamanetwork.com/journals/jama/article-abstract/2767353
44. Datasheets for datasets. (n.d.). Accessed January 21, 2024. https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/abstract
45. Health equity intervention and action principles. (n.d.). Accessed January 22, 2024. https://www.cdc.gov/healthequity/whatis/healthequityinaction/topics/he-intervention-action-principles.html
46. Dankwa-Mullan I, Scheufele E, Matheny ME, et al. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. *J Health Care Poor Underserved*. 2021;32(2S):300-317. https://doi.org/10.1353/hpu.2021.0065.
47. What is explainable AI? (n.d.). Accessed January 19, 2024. https://www.ibm.com/topics/explainable-ai#:~:text=Explainable%20artificial%20intelligence%20(XAI)%20is,expected%20impact%20and%20potential%20biases
48. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. *NPJ Digit Med*. 2020;3(1):41. https://doi.org/10.1038/s41746-020-0253-3
49. Liu Y, Chen P-HC, Krause J, Peng L. How to read articles that use machine learning: users' guides to the medical literature. *JAMA*. 2019;322(18):1806-1816. https://doi.org/10.1001/jama.2019.16489
50. MedWatch: the FDA safety information and adverse event reporting program. (n.d.). Accessed December 9, 2023. https://www.fda.gov/safety/medwatch-fda-safety-information-and-adverse-event-reporting-program#
51. Potnis KC, Ross JS, Aneja S, Gross CP, Richman IB. Artificial intelligence in breast cancer screening: evaluation of FDA device regulation and future recommendations. *JAMA Intern Med*. 2022;182(12):1306-1312. https://doi.org/10.1001/jamainternmed.2022.4969
52. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med*. 2021;27(4):582-584. https://doi.org/10.1038/s41591-021-01312-x
53. MAUDE—manufacturer and user facility device experience. Accessed January 31, 2024. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm
54. Russell RG, Novak LL, Patel M, et al. Competencies for the use of artificial intelligence-based tools by healthcare professionals. *Acad Med*. 2023;98(3):348-356. https://doi.org/10.1097/ACM.0000000000004963. Selected as June 2023 Johns Hopkins Med Ed Team 'Must Read.'
55. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA*. 2019;322(24):2377-2378. https://doi.org/10.1001/jama.2019.18058
56. Artificial intelligence and machine learning in software as a medical device. (n.d.). Accessed January 10, 2024. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device
57. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD). (n.d.). Accessed December 28, 2023. https://www.fda.gov/media/122535/download
58. Medicare and Medicaid Programs: electronic health record incentive program – stage 2. (n.d.). Accessed January 20, 2024. https://www.federalregister.gov/documents/2012/09/04/2012-21050/medicare-and-medicaid-programs-electronic-health-record-incentive-program-stage-2
59. Greenes RA, Del Fiol G., eds. *Clinical Decision Support and Beyond: Progress and Opportunities in Knowledge-Enhanced Health and Healthcare*. 3rd ed. Academic Press; 2023.
60. HL7 International. Accessed September 20, 2024. https://www.hl7.org
61. SNOMED. Accessed September 20, 2024. https://www.snomed.org/
62. Finnell JT, Dixon BE, eds. *Clinical Informatics Study Guide: Text and Review*. Springer International Publishing; 2022. https://doi.org/10.1007/978-3-030-93765-2
63. Bates DW, Auerbach A, Schulam P, Wright A, Saria S. Reporting and implementing interventions involving machine learning and artificial intelligence. *Ann Intern Med*. 2020;172(11_Suppl):S137-S144. https://doi.org/10.7326/M19-0872
64. Carr NG. *The Glass Cage: How Our Computers Are Changing Us*. New York, London: Norton & Company, 2015.
65. AI in healthcare: risk management, trust, and liability. Accessed 5 January, 2024. https://www.dcinetwork.org/events/190
66. Bates DW, Levine DM, Salmasian H, et al. The safety of inpatient health care. *N Engl J Med*. 2023;388(2):142-153.
67. European Union. The EU Artificial Intelligence Act. Accessed December 5, 2024. https://artificialintelligenceact.eu
68. Whitehouse. Blueprint for an AI Bill of Rights. Accessed September 20, 2024. https://www.whitehouse.gov/ostp/ai-bill-of-rights/

Journal of the American Medical Informatics Association, 2024, 31, 2730–2739
https://doi.org/10.1093/jamia/ocae209
Case Report