ConnectedHealthInitiative

September 29, 2022

Mr. Mark Przybocki
U.S. National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, Maryland 20899

RE: Comments of the Connected Health Initiative to the National Institute of Standards and Technology on its Second Draft of the Artificial Intelligence Risk Management Framework

The Connected Health Initiative appreciates the opportunity to submit views to the National Institute of Standards and Technology (NIST) on its latest draft of the Artificial Intelligence (AI) Risk Management Framework (RMF).[1]

CHI is the leading effort by stakeholders across the connected health ecosystem to clarify outdated health regulations, encourage the use of remote monitoring (RM), and support an environment in which patients and consumers can see improvement in their health. This coalition of leading mobile health companies and stakeholders urges Congress, the Office of the National Coordinator for Health Information Technology (ONC), the Food and Drug Administration (FDA), the Centers for Medicare & Medicaid Services (CMS), and other regulators, policymakers, and researchers to adopt frameworks that encourage mobile health innovation using interoperable data while keeping sensitive health data private and secure. CHI is a longtime supporter of ONC in its efforts to establish rules prohibiting illegal information blocking, which are critical to realizing a connected care continuum. NIST's planned voluntary AI RMF—and the efforts of numerous agencies with respect to AI policy and regulation—directly impact the CHI community. We support NIST's goal of helping designers, developers, users, and evaluators of AI systems better manage risks across the AI lifecycle.

Artificial/augmented intelligence (AI) and machine learning (ML), powered by streams of data and advanced algorithms, have incredible potential to improve healthcare, prevent hospitalizations, reduce complications, and increase patient engagement. Yet, applications of AI in healthcare have also given rise to a variety of potential challenges for policymakers to consider, including quality assurance, adaptiveness, ethics, oversight, notice/consent, and data bias. The FDA must take a leading role in responsibly bringing AI medical devices to the marketplace, and we support FDA's continued leadership to develop a governance framework for AI meeting the definition of a medical device under the Federal Food, Drug, and Cosmetic Act (FD&C Act).

[1] https://www.nist.gov/itl/ai-risk-management-framework.

As part of its commitment to responsibly advance AI in healthcare, CHI has assembled a Health AI Task Force consisting of a range of innovators and thought leaders. CHI's AI Task Force has developed a range of resources that we urge NIST to align its AI RMF with, which are also appended to this comment letter:

- **Health AI Policy Principles:** https://bit.ly/3m9ZBLv
- **Why AI? Considerations for Use of Artificial Intelligence in States' Medicaid and CHIP Programs:** https://bit.ly/2Y2FJle
- **Good Machine Learning Practices for FDA-Regulated AI:** https://bit.ly/2YaYljk
- **Advancing Transparency for Artificial Intelligence in the Healthcare Ecosystem:** https://bit.ly/3n36WO5

Further, the CHI's AI Task Force continues to examine areas where the development of cross-digital health ecosystem consensus would assist policymakers at all levels and expects to develop further deliverables in areas including data bias mitigation and best practices for the use of real-world evidence.

Noting our general support for NIST's efforts to develop a voluntary prioritized, flexible, risk-based, outcome-focused, and cost-effective AI RMF, we offer the following input on the latest draft of the AI RMF:

- The AI RMF latest draft generally covers and addresses AI risks, while maintaining a scalable approach to risk management, similar to the approach taken in NIST's Cybersecurity Framework. Such an approach recognizes that appropriate risk management practices reflect specific fact patterns and the unique risks posed by certain uses of AI. The latest draft of the AI RMF reflects that trustworthiness is sector-dependent and that particular risk management practices develop trust. Moreover, the latest draft highlights that risk management processes assist organizations to design and develop trustworthy solutions. We generally support NIST's proposed functions, categories, and subcategories that will assist organizations in using the NIST AI RMF once finalized.

  We also support the AI RMF latest draft's efforts to map to, and rely on, international standards such as from ISO-IEC/JTC-1-SC 42, IEEE, ASTM, and SAE. We recognize, however, that international standardization for AI risk management is not yet robust. As AI standardization matures, it will be important that NIST continue to update its AI RMF periodically (ideally annually), and that its collaborators abroad do the same.

- We support maintaining the latest draft of the AI RMF's reinforcement that it is voluntary, much like the NIST Cybersecurity Framework. NIST should recognize that some may attempt to position the NIST AI RMF as a mandatory standard of behavior (e.g., in litigation), and we encourage NIST's AI RMF to directly address this concern and reinforce in the NIST AI RMF that (1) adoption of the RMF is voluntary for the private sector and (2) that the RMF is not, and is not intended to be offered as, a baseline for behavior norms in the context of litigation.

- We again encourage NIST to ensure that the RMF prioritizes the design of AI systems, from the earliest phase of product development, being informed by real-world workflows, human-centered design and usability principles, and end-user needs. We recognize that this concept is reflected in the latest draft of the AI RMF, but to specifically encourage NIST to call on AI developers to build in AI quality through the data streams mentioned above "by design." AI systems solutions should facilitate a transition to changes in the delivery of goods and services that benefit consumers and businesses. The design, development, and success of AI should take advantage of collaboration and dialogue among users, AI technology developers, and other interested parties in order to have all perspectives reflected in AI solutions. As this concept must run across sectors and AI use cases, we call on NIST to advance thoughtful design principles in the RMF, and label them as such.

  We appreciate the latest draft of the AI RMF's enabling the needs of different sectors and use cases to be addressed modularly. Different industries and sectors will need to focus on other impacted stakeholder groups due to, among other requirements, regulatory requirements (as a further example, the automobile industry may need to assess risks to the vehicle driver, to other drivers on the road, or to pedestrians), which underscores that the AI RMF must not impede different sectors' unique risk management needs.

  The required depth for risk analysis will also depend on established precedent. Using the healthcare sector as an example, for lower risk devices (e.g., a digital thermometer), showing the device is useable and useful to the intended user is sufficient; for higher risk devices (e.g., a pacemaker), manufacturers may be asked to show a more comprehensive risk/benefit analysis in the intended context of use, or a suitable clinical study. This is to say that, even within sectors, risk analysis will depend heavily upon the purpose the device serves—another indicator that different sectors will need flexibility in risk management processes.

- CHI appreciates that NIST appropriately addresses the need to combat harmful biases in AI datasets. Our community is working to develop a consensus standard on how to validate that biases are being identified and appropriately mitigated, and to establish an adequate infrastructure of test beds for making such standards operational. We welcome the AI risk management framework as a driver of partnership between technology developers and the government, and others, to address how to make AI data sets appropriately representative of the populations/communities AI tools are intended to serve and benefit.

  The success of AI depends on ethical development and use, which in turn impacts user trust and use of AI. The RMF should promote existing and emerging ethical norms for broader adherence by AI technologists, innovators, computer scientists, and those who use such systems. We support NIST providing for stakeholders' approaches to AI to duly consider ethics to:

  - Ensure that AI solutions align with all relevant ethical obligations, across the lifecycle of algorithms or models.
  - Encourage the development or updates of ethical guidelines to address issues emerging with the AI's use, as needed.
  - Maintain consistency with international conventions on human rights.
  - Ensure that AI is inclusive such that AI solutions beneficial to consumers are developed across socioeconomic, age, sex, gender, geographic origin, and other groupings.
  - Reflect that AI tools may reveal extremely sensitive and private information about a user or group and ensure that laws protect such information from being used to discriminate against certain consumers.

- CHI supports NIST's efforts in the latest draft of the AI RMF in helping providers, technology developers and vendors, and others involved to understand the distribution of risk and liability in building, testing, deploying, and even decommissioning AI tools. The RMF should advance the appropriate distribution and mitigation of AI-related risks and liabilities. That is, those in the value chain with the ability to minimize risks based on their knowledge and ability to mitigate should be incentivized take reasonable steps to do so. Already, other U.S. government agencies are seeking to advance improper and infeasible approaches to AI liability,[2] which NIST should take steps to avoid, both through updates to its AI RMF itself as well as through cross-agency outreach.

  Further, the RMF should clearly state that those developing, offering, or testing AI systems provide truthful and easy to understand representations regarding intended use and risks that would be reasonably understood by those impacted (as well as expected to be impacted) to use the AI solution. References to "intended use" and "context of use" as concepts is an appropriate approach for AI risk management, which encourages AI developers to identify and model to address challenges prior to algorithm development, which in turn allows developers to identify and evaluate potential threats more effectively.

- CHI appreciates NIST's commitment in the latest draft to have the AI RMF support organizations' abilities to operate under applicable domestic and international legal or regulatory regimes. We urge NIST to include a priority for

[2] *E.g.*, https://www.federalregister.gov/documents/2022/08/04/2022-16217/nondiscrimination-in-health-programs-and-activities.

aligning, where appropriate, with international efforts, and upon completion, promoting the NIST AI RMF for use internationally. Already, developers of AI face top-down and one-size-fits-all mandates that substantially impede their ability to develop and utilize AI across a range of use cases. It is crucial that the NIST AI RMF be offered as an alternative to such mandates, or at least have a positive influence on mandates, from other jurisdictions.

CHI appreciates NIST's consideration of the above views. AI offers immense potential for widespread societal benefit, which is why NIST's voluntary RMF should foster investment and innovation in any way practicable.

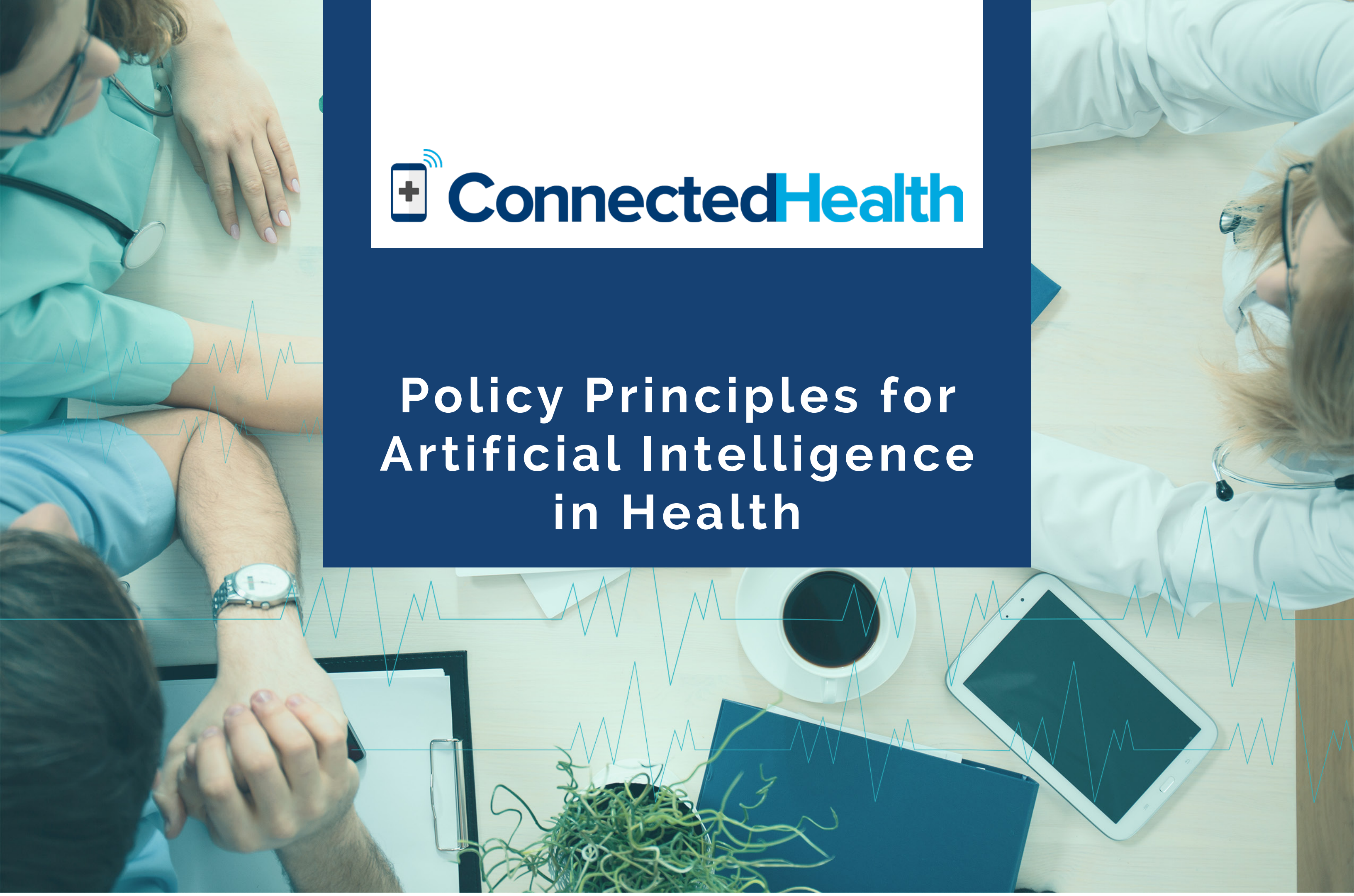
We urge NIST to contact the undersigned with any questions or ways that we can assist moving forward.

Sincerely,

Brian Scarpelli
Senior Global Policy Counsel

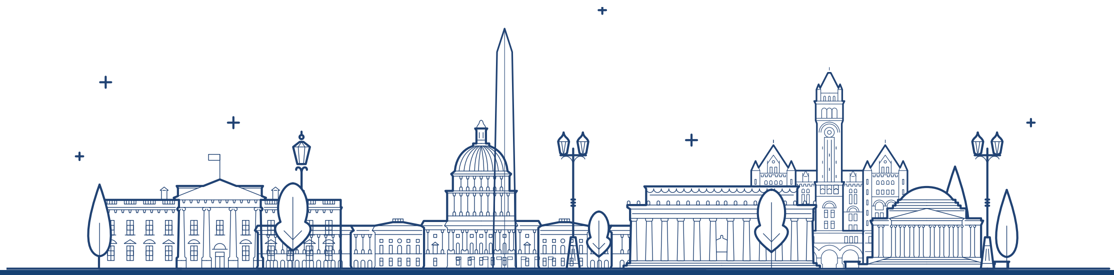
Leanna Wade
Policy Associate

Connected Health Initiative
1401 K St NW (Ste 501)
Washington, DC 20005
202-331-2130



ConnectedHealth

Policy Principles for Artificial Intelligence in Health



Connected Health is an initiative
of ACT | The App Association

1401 K Street NW Suite 501
Washington, DC 20005

202.331.2130

connectedhi.com

#connectedhealth

/ConnectedHealthInitiative

Policy Principles for AI in Health

Today, there are already many examples of AI systems, powered by streams of data and advanced algorithms, improving healthcare by preventing hospitalizations, reducing complications, decreasing administrative burdens, and improving patient engagement. AI systems offer the promise to rapidly accelerate and scale such results and drive a fundamental transformation of the current disease-based system to one that supports prevention and health maintenance. Nonetheless, AI in healthcare has the potential to raise a variety of unique considerations for U.S. policymakers.

Many organizations are taking steps to proactively address adoption and integration of AI into health care and how it should be approached by clinicians, technologists, patients and consumers, policymakers, and other stakeholders, such as the Partnership for AI, Xavier Health, the American Medical Association, and the Association for the Advancement of Medical Instrumentation and BSI. Building on these important efforts, the Connected Health Initiative's (CHI) Health AI Task Force is taking the next step to address the role of AI in healthcare.

First, AI systems deployed in healthcare must advance the "quadruple aim" by improving population health; improving patient health outcomes and satisfaction; increasing value by lowering overall costs; and improving clinician and healthcare team well-being. Second, AI systems should:

- Enhance access to health care.
- Empower patients and consumers to manage and optimize their health.
- Facilitate and strengthen the relationship and communication that individuals have with their health care team.
- Reduce administrative and cognitive burdens for patients and their health care team.

To guide policymakers, we recommend the following principles to guide action:

- **National Health AI Strategy:** Many of the policy issues raised below involve significant work and changes that will impact a range of stakeholders. The cultural, workforce training and education, data access, and technology-related changes will require strong guidance and coordination. Given the significant role of the government in the regulation, delivery, and payment of healthcare, as well as its role as steward of significant amounts of patient data, a federal healthcare AI strategy incorporating guidance on the issues below will be vital to achieving the promise that AI offers to patients and the healthcare sector. Other countries have begun to take similar steps (e.g., The UK's Initial Code of Conduct for Data Driven Care and Technology) and it is critical that U.S. policymakers collaborate with provider organizations, other civil society organizations, and private sector stakeholders to begin similar work.

- **Research:** Policy frameworks should support and facilitate research and development of AI in healthcare by prioritizing and providing sufficient funding while also ensuring adequate incentives (e.g., streamlined availability of data to developers, tax credits) are in place to encourage private and non-profit sector research. Clinical validation and transparency research should be prioritized and involve collaboration among all affected stakeholders who must responsibly address the ethical, social, economic, and legal implications that may result from AI applications in healthcare. Further, public funding and incentives should be conditioned on promoting the medical commons in order to advance shared knowledge, access, and innovation.
- **Quality Assurance and Oversight:** Policy frameworks should utilize risk-based approaches to ensure that the use of AI in healthcare aligns with recognized standards of safety, efficacy, and equity. Providers, technology developers and vendors, health systems, insurers, and other stakeholders all benefit from understanding the distribution of risk and liability in building, testing, and using healthcare AI tools. Policy frameworks addressing liability should ensure the appropriate distribution and mitigation of risk and liability. Specifically, those in the value chain with the ability to minimize risks based on their knowledge and ability to mitigate should have appropriate incentives to do so. Some recommended guidelines include:
  - Ensuring AI in healthcare is safe, efficacious, and equitable.
  - Ensuring algorithms, datasets, and decisions are auditable and when applied to medical care (such as screening, diagnosis, or treatment) are clinically validated and explainable.
  - AI developers should consistently utilize rigorous procedures and must be able to document their methods and results.
  - Those developing, offering, or testing healthcare AI systems should be required to provide truthful and easy to understand representations regarding intended use and risks that would be reasonably understood by those intended, as well as expected, to use the AI solution.
  - Adverse events should be timely reported to relevant oversight bodies for appropriate investigation and action.

- **Thoughtful Design:** Policy frameworks should require design of AI systems in health care that are informed by real-world workflow, human-centered design and usability principles, and end-user needs. Also, AI systems should help patients, providers, and other care team members overcome the current fragmentation and dysfunctions of the healthcare system. AI systems solutions should facilitate a transition to changes in care delivery that advance the quadruple aim. The design, development, and success of AI in healthcare should leverage collaboration and dialogue between caregivers, AI technology developers, and other healthcare stakeholders in order to have all perspectives reflected in AI solutions.

- **Access and Affordability:** Policy frameworks should ensure AI systems in health care are accessible and affordable. Significant resources may be required to scale systems in health care and policy-makers must take steps to remedy the uneven distribution of resources and access. There are varied applications of AI systems in health care such as research, health administration and operations, population health, practice delivery improvement, and direct clinical care. Payment and incentive policies must be in place to invest in building infrastructure, preparing personnel and training, as well as developing, validating, and maintaining AI system with an eye toward ensuring value. While AI systems should help transition to value-based delivery models by providing essential population health tools and providing enhanced scalability and patient support, in the interim payment policies must incentivize a pathway for the voluntary adoption and integration of AI systems into clinical practice as well as other applications under existing payment models.

- **Ethics:** Given the longstanding, deeply rooted, and well-developed body of medical and biomedical ethics, it will be critical to promote many of the existing and emerging ethical norms of the medical community for broader adherence by technologists, innovators, computer scientists, and those who use such systems. Healthcare AI will only succeed if it is used ethically to protect patients and consumers. Policy frameworks should:

  - Ensure AI in healthcare is safe, efficacious, and equitable.

  - Ensure that healthcare AI solutions align with all relevant ethical obligations, from design to development to use.

  - Encourage the development of new ethical guidelines to address emerging issues with the use of AI in healthcare, as needed.

  - Ensure consistency with international conventions on human rights.

  - Ensure that AI for health is inclusive such that AI solutions beneficial to patients are developed across socioeconomic, age, gender, geographic origin, and other groupings.

  - Reflect that AI for health tools may reveal extremely sensitive and private information about a patient and ensure that laws protect such information from being used to discriminate against patients.

- **Modernized Privacy and Security Frameworks:** While the types of data items analyzed by AI and other technologies are not new, this analysis provides greater potential utility of those data items to other individuals, entities, and machines. Thus, there are many new uses for, and ways to analyze, the collected data. This raises privacy issues and questions surrounding consent to use data in a particular way (e.g., research, commercial product/service development). It also offers the potential for more powerful and granular access controls for patients. Accordingly, any policy framework should address the topics of privacy, consent, and modern technological capabilities as a part of the policy development process. Policy frameworks must be scalable and assure that an individual's health information is properly protected, while also allowing the flow of health information. This information is necessary to provide and promote high-quality healthcare and to protect the public's health and well-being. There are specific uses of data that require additional policy safeguards, i.e., genomic information. Given that one individual's DNA includes potentially identifying information about even distant relatives of that individual, a separate and more detailed approach may be necessary for genomic privacy. Further, enhanced protection from discrimination based on pre-existing conditions or genomic information may be needed for patients. Finally, with proper protections in place, policy frameworks should also promote data access, including open access to appropriate machine-readable public data, development of a culture of securely sharing data with external partners, and explicit communication of allowable use with periodic review of informed consent.

- **Collaboration and Interoperability:** Policy frameworks should enable eased data access and use through creating a culture of cooperation, trust, and openness among policymakers, health AI technology developers and users, and the public.

- **Workforce Issues and AI in Healthcare:** The United States faces significant demands on the healthcare system and safety net programs due to an aging population and a wave of retirements among practicing care workers. And lower birth rates mean that fewer young people are entering the workforce. Successful creation and deployment of AI-enabled technologies which help care providers meet the needs of all patients will be an essential part of addressing this projected shortage of care workers. Policymakers and stakeholders will need to work together to create the appropriate balance between human care and decision-making and augmented capabilities from AI-enabled technologies and tools.

- **Bias:** The bias inherent in all data as well as errors will remain one of the more pressing issues with AI systems that utilize machine learning techniques in particular. In developing and using healthcare AI solutions, these data provenance and bias issues must be addressed. Policy frameworks should:

  - Require the identification, disclosure, and mitigation of bias while encouraging access to databases and promoting inclusion and diversity.

  - Ensure that data bias does not cause harm to patients or consumers.

- **Education:** Policy frameworks should support education for the advancement of AI in healthcare, promote examples that demonstrate the success of AI in healthcare, and encourage stakeholder engagements to keep frameworks responsive to emerging opportunities and challenges.
  - Patients and consumers should be educated as to the use of AI in the care they are receiving.
  - Academic/medical education should include curriculum that will advance health care providers' understanding of and ability to use health AI solutions. Ongoing continuing education should also advance understanding of the safe and effective use of AI in healthcare delivery.

ConnectedHealth

Why AI? Considerations for Use of Artificial Intelligence in States' Medicaid and CHIP Programs

June 23, 2020

I. Executive Summary

Artificial/augmented intelligence (AI) has the potential to transform the healthcare system and the radically improve the experiences of patients and healthcare providers throughout America's healthcare ecosystem. Experts believe that the responsible use of AI will enhance the quality of care, prevent hospitalizations, reduce complications, and improve patient engagement while also lowering costs. At the same time, AI healthcare applications will also give rise to a variety of potential challenges for policymakers to consider including concerns about privacy, bias, inclusion, and transparency.

As administrators of two of the largest healthcare providers in America, Medicaid programs and the Children's Health Insurance Program (CHIP), state governments have an important role in defining the future of AI in American healthcare. They have an opportunity to not only leverage AI in service of their states most vulnerable citizens, but also ensure AI is adopted in the most responsible, transparent and equitable way possible.

Responsible use of AI has the potential to dramatically reduce administrative burdens and free up Medicaid and CHIP programs resources to focus on improving care for patients and permitting resource redeployment to better serve the most vulnerable populations and to mitigate and eliminate disparities in healthcare at all levels. Furthermore, AI has also demonstrated an ability to help manage public health emergencies (e.g., COVID-19) at the state level and to aid in related recovery efforts.

To help state policymakers navigate the opportunities and challenges of adopting AI-based technologies, the Connected Health Initiative convened a multidisciplinary group of experts to to develop a robust set of AI Policy Principles and produce the following analysis of the key considerations state policymakers should keep in mind as they consider AI's proper role in their state's Medicaid and CHIP programs.

At the core of our recommendations is the use of AI tools to advance the "Quadruple Aim," a widely accepted compass for optimizing the performance of health systems based on the work of the Institute for Healthcare Improvement. These goals include:

1. enhancing population health;
2. improving patient experience, satisfaction, and health outcomes;
3. better clinician and healthcare team experience and satisfaction;
4. lowering overall costs of healthcare.

CHI recommends that states develop an integrated and robust state health IT strategy to maximize the effectiveness of these technologies for their citizens. This strategy should be developed in partnership with a wide range of stakeholders from within healthcare industry, advocacy organizations, academia, patient organizations and beyond. Some of the core areas this strategy should address are: research and development, quality assurance and oversight, human-centered design, access and affordability, ethics, privacy and security, interoperability, workforce development, and education.

II. Introduction/About the CHI Health AI Task Force

A. About CHI and its Health AI Task Force

The Connected Health Initiative (CHI) is the leading multistakeholder advocacy organization for the responsible deployment and use of innovative connected technologies in the delivery of healthcare. We represent a broad and diverse consensus of healthcare stakeholders from physician groups to technology companies. CHI seeks policy and legal changes that will allow all Americans to realize the benefits of an information and communications technology-enabled American healthcare system. CHI is committed to advancing technology use across healthcare use cases that will improve outcomes for all patients as well as mitigate and eliminate disparities in healthcare at all levels. For more information, see www.connectedhi.com.

Artificial/augmented intelligence (AI) has the potential to positively transform the healthcare system and the experiences of patients and healthcare providers throughout the healthcare ecosystem. AI has incredible potential to improve healthcare, prevent hospitalizations, reduce complications, and improve patient engagement. Along with this promise, CHI recognizes that AI healthcare applications also give rise to a variety of potential challenges for policymakers to consider, including notice/consent, bias, inclusion, and transparency, among others.

As healthcare AI innovations advance and begin navigating the regulatory processes on the way to market, policymakers at both the legislative and regulatory levels are considering whether there is a need for accompanying policy changes.

Recognizing the significant role that states play through their administration of Medicaid programs and the Children's Health Insurance Program (CHIP), CHI has prepared this whitepaper to aide policymakers in their policymaking processes. CHI's Health AI Task Force—consisting of a range of subject matter expert stakeholders from throughout the healthcare continuum—developed a robust set of AI Policy Principles[i] which provide a foundation our analysis of considerations that state policymakers should keep in mind as they consider AI's role in their state's Medicaid and CHIP programs.

B. Overview of AI

Although AI has various definitions based on context and sector-specific qualifiers, most individuals in the field would agree that AI includes systems or machines that mimic human intelligence to perform tasks.[ii] AI is an evolving constellation of technologies that enables computers to simulate elements of human thinking – learning and reasoning among them. Furthermore, AI is a multidimensional term that encompasses a range of approaches and technologies, such as machine learning (ML) and deep learning, where an algorithm can adapt by “learning” when exposed to new inputs, allowing for independent or assisted decision making.

AI-driven algorithmic decision tools and predictive analytics are having, and will continue to have, substantial direct and indirect effects on Americans. Some forms of AI are already in use to improve American consumers’ lives today. For example, AI can augment efforts to detect financial and identity theft and to protect the communications networks upon which Americans rely against cybersecurity threats.

Breakthroughs are expected to create a $126 billion AI marketplace by 2025 with the opportunity for far-reaching benefits.[iii] If policymakers navigate the challenges and opportunities effectively, AI will improve American consumers’ lives through faster and better-informed decision making enabled by cutting-edge distributed cloud computing. AI will also provide for more effective governance through its ability to enhance infrastructure foresight and support efficient budgeting decisions. AI will beneficially impact every aspect of Americans’ lives if we encourage ethical innovation at AI’s beginning stages.

Today, Americans encounter AI in their lives incrementally through improvements in computer-based services in the form of streamlined processes, image analysis, and voice recognition. It is evident today that this “narrow” AI approach is already providing significant societal benefits. For example, AI-driven software products and services enabled countless disabled Americans to experience sensations that people without a disability can perceive on a day-to-day basis, revolutionizing and improving their day-to-day lives.[iv]

Along with these transformative benefits, AI raises a variety of unique considerations for societal concerns that policymakers must address to realize the promise of AI. Policymakers must find a balanced approach to the implementation of AI innovation with necessary safeguards to protect consumers and society. It is important that policymakers consider the variety of stakeholders that AI may influence. This is especially true in the healthcare context when making statutory and regulatory changes impacting AI. Such changes must be based on risk of harm and benefit accounting for a host of factors, including evidence of safety, efficacy and equity including addressing bias; AI system methods, level of automation, transparency and conditions of deployment. Given the demonstrated benefits of AI across numerous consumer and enterprise use cases, state policymakers should strive to support AI systems that advance the quadruple aim and make these benefits available to their citizens across the healthcare spectrum, particularly for Medicaid and CHIP beneficiaries.

III. Why Do Medicaid and CHIP Need AI?

Medicaid provides health coverage to millions of low-income Americans and is one of the largest payers for healthcare within the United States.[v] Currently, 63.9 million people are enrolled for Medicaid services including low-income adults, children, pregnant women, elderly adults, and individuals with disabilities.[vi]

CHIP programs provide health coverage to eligible children, through both Medicaid and separate CHIP programs, and is administered by states, according to federal requirements.[vii] Approximately 9.6 million children are enrolled in CHIP.[viii]

States administer Medicaid and CHIP funds, according to federal requirements, and both state and the federal governments contribute funding to Medicaid and CHIP programs on an annual basis. While these funds serve millions of Americans, state policymakers need to think about the future and how their state can continue to serve each state's most vulnerable populations despite limited budgets and resources. Therefore, it is important that states consider responsibly incorporating new and innovative technologies such as AI into their everyday work. AI can dramatically reduce administrative burdens, improve physicians' ability to care for their patients, and permit resource redeployment within Medicaid systems and CHIP programs to better serve the most vulnerable populations.

Furthermore, AI has also demonstrated an ability to help manage public health emergencies at the state level. In addressing the COVID-19 pandemic, health authorities found that AI greatly assists in population health management (infection trends, resource management, etc.), as well as in diagnosis and treatment of individuals.[ix] Additionally, AI has played a role in tracking helpful research that will contribute to a potential vaccine for the COVID-19 virus.[x]

We urge state policymakers, when considering the value of AI in healthcare, to view the proposition through the lens of the "quadruple aim" framework. Built on the Institute for Healthcare Improvement's "triple aim,"[xi] a widely accepted compass to optimize health system performance,[xii] the quadruple aim focuses on four key metrics for optimizing health systems to meet the needs a wide range of key stakeholders and communities. The four areas are (1) enhancing population health; (2) improving patient experience, satisfaction, and health outcomes; (3) better clinician and healthcare team experience and satisfaction; and (4) lowered overall costs of healthcare.

Further, across the country, disparities in healthcare are sizable and growing, caused by barriers that exist at all levels, exacerbated by the ongoing COVID-19 public health emergency.[xiii] To address these disparities and achieve health equity, state policymakers should identify potential bias in data collection and responsibly utilize AI tools. Great strides can be taken to achieve health equity (and aid in a lasting recovery) through, for example, the collection and use of health and/or SDOH data disaggregated by race, ethnicity, gender, disability, and other characteristics, consistent with the reocmmendations below.

A. Improving population health management

Population health[xiv] management is an essential ingredient to improve overall health outcomes and arrest rising health care costs. Population health management involves aggregation and analysis of huge amounts of data from divergent sources, something that can be potentially streamlined through robust and powerful AI systems. AI-powered tools can responsibly collect patient generated health data and deliver clinically-backed interventions to treat patients where they are.

AI-enabled tools offer great promise in overcoming the challenges faced by clinicians, health systems, health plans, and public health officials working to advance population health management and public health. Social determinants of health (SDOH) – social factors as diverse as income, access to transportation and healthy food, and education – can also provide key indicators of health and well-being, helping providers and health plans manage population health. This can provide public health officials, healthcare systems, and providers near real time access to essential and actionable data to assist with more timely and accurate population level disease surveillance and assessments of disparities and health care resource distribution.

As more systems are created and deployed, the opportunity for AI to help improve healthcare outcomes across communities is significant, with estimates suggesting outcomes could be improved by 30-40 percent.[xv]

B. Improving patient experience, satisfaction, and outcomes

One of the more significant critiques of healthcare systems around the world is that they fail in many respects to meet patients' expectations around access to care, ease of use, and care continuity and coordination.

All too often, patients must make multiple visits, shuffling between a general practitioner and a specialist. AI-enabled tools can reduce paperwork burdens, center care around the location of the patient and enhance the ability to manage and understand how to sustain health or manage a disease. AI systems can also provide patients and their health care teams with timely, essential information, and ongoing support that is not currently available.

Given the unique and diverse needs of Medicaid and CHIP beneficiaries, the vast majority of whom lack access to other affordable health insurance and have limited ability to pay out-of-pocket costs for acute or long-term care, AI systems will be essential for human caregivers and clinicians to extend their reach and coverage (e.g., creating efficiencies, highlighting relevant information, and presenting gaps in care that beneficiaries receive) of this vulnerable population of patients efficiently and in as tailored a fashion as possible.

C. Improving clinician and healthcare team experience and satisfaction

Clinicians and extended healthcare teams are experiencing record levels of burn-out and dissatisfaction which is largely attributable to growing demands of administrative paperwork coupled with compounding rates of new medical knowledge and data generation. Deployment of AI-enabled tools can drastically improve clinician and healthcare team satisfaction using tools that help them more efficiently screen, diagnose, treat, and monitor patients and remove and/or reduce time-consuming (and often mundane) tasks.

D. Reducing healthcare costs

States continue to struggle with both rising and absolute costs of providing healthcare to their citizens. Nationally, health spending is projected to grow at an average rate of 5.5 percent per year for 2018-27, reaching nearly $6 trillion by 2027.[xvi] Further, according to the Association of American Medical Colleges (AAMC), there will be a deficit of skilled healthcare professionals to serve our population with a rising average age and life expectancy. AAMC's report indicates that by 2032, the United States will face a shortage of 46,900 to 121,900 physicians leaving the United States with an unsustainable healthcare shortage unless something is done.[xvii]

AI is a critical component to resolving these rapidly approaching healthcare concerns. Implementation of AI healthcare tools can not only reduce overall healthcare costs directly, but also contribute to increased efficiencies that address challenges such as lack of care coordination, overtreatment, low value of care, burdensome administrative processes, and identification of fraud and abuse within medical systems. These efficiencies will enable professional medical staff to spend more time with patients by utilizing tools that rely on AI to analyze large datasets, facilitating more informed patient care.

Healthcare experts see enormous promise in AI's ability to more accurately capture and leverage the range of health data available. Estimates suggest successful use of AI applications will create $150 billion in annual savings for the United States healthcare economy alone by 2026.[xviii] We note that this savings estimate should be considered conservative, as it only includes a "top 10" of AI scenarios, such as assisted surgery, virtual nursing assistants, and administrative workflow assistance. More efficient and timely use of health data will provide many further benefits across a range of further scenarios and use cases. Because improved patient outcomes for Medicaid and CHIP beneficiaries will entail allotting resources to services other than those addressing acute and chronic illnesses, AI can help Medicaid bring the right resources to the right areas to support additional services such as therapy, tailored case management, habilitative services, and transport and translation costs.

Further, healthcare administrative costs (e.g., billing) are a continuing challenge that cannot be understated. The administrative costs of the U. S. healthcare system are estimated to be 31 percent of total healthcare expenditures.[xix] Administrative AI's potential to help address spiraling costs in healthcare is already being realized today.

IV. Responsibly Implementing AI in Medicaid and CHIP Programs

Not only do state governments play a critical role in the regulation, delivery, and payment of healthcare, but they are also stewards of significant amounts of patient data. State policymakers have a responsibility to ensure that AI systems are effectively and responsibly implemented within Medicaid and CHIP programs and address specific issues in coordination with key stakeholders. To aid in this process, CHI has developed this framework for action:

State Health AI Strategy: State governments need a coordinated strategy to deliver on the full promise of AI to the healthcare sector and the patients it serves. For example, navigating the cultural, workforce training and education, data access, and technology-related changes will require strong guidance and coordination. Implementing AI within the healthcare industry also will impact a wide range of stakeholders, and it is critical that policymakers collaborate with provider organizations, other civil society organizations, and private sector stakeholders in the development of such strategies.

Research: State policy frameworks should support and facilitate research and development of AI in healthcare by prioritizing and providing sufficient funding while also ensuring adequate incentives (e.g., streamlined availability of data to developers, tax credits) are in place to encourage private and non-profit sector research. Clinical validation and transparency research should be prioritized and involve collaboration among all affected stakeholders who must responsibly address the ethical, social, economic, and legal implications that may result from AI applications in healthcare. Further, public funding and incentives should be conditioned on promoting the medical commons in order to advance shared knowledge, access, and innovation.

Quality Assurance and Oversight: Policy frameworks, within states' authorities, should utilize risk-based approaches to ensure that the use of AI aligns with recognized standards of safety, efficacy, and equity. CHI recommends state policymakers ensure that:

- AI in healthcare is safe, efficacious, and equitable. AI should utilize risk-based security practices to ensure health data privacy at all points in the care continuum.
- Algorithms, datasets, and decisions are auditable and when applied to medical care (such as screening, diagnosis, or treatment) are clinically validated and reasonably understood/explained.
- AI developers should consistently utilize rigorous procedures for development, testing, and validation and must be able to document their methods and results.
- Those developing, offering, or testing healthcare AI systems should be required to provide truthful and easy to understand representations regarding intended use and risks that would be reasonably understood by those intended to use the AI solution.
- Adverse events should be reported in a timely manner to relevant oversight bodies for appropriate investigation and action.

In addition, policymakers should explore using AI tools to improve overruse and fraud detection in the context of Medicaid and CHIP patient care and reimbursement through identification of anomalies and trends.

Distribution of Liability: Providers, technology developers and vendors, health systems, insurers, and other stakeholders all benefit from understanding the distribution of risk and liability in building, testing, and using healthcare AI tools. Policy frameworks addressing liability should ensure the appropriate distribution and mitigation of risk and liability. Specifically, those in the value chain with the ability to minimize risks based on their knowledge and ability to mitigate should have appropriate incentives to do so.

Human-Centered Design: Policy frameworks should require AI systems adopted in healthcare systems to be developed and tested using human-centered design and usability principles. This should include examination of real-world workflow and beneficiary needs – particularly those with disabilities and others with unique needs specific to certain Medicaid and CHIP beneficiary populations. The design and development of AI in healthcare should also leverage collaboration and dialogue between caregivers, AI technology developers, and other healthcare stakeholders in order to have all perspectives reflected in AI solutions. Effectively designed AI systems can help patients, providers, and other care team members overcome the current fragmentation and dysfunctions of the healthcare system and facilitate changes in care delivery that advance the quadruple aim.

Access and Affordability: Policy frameworks should ensure the us of AI systems in healthcare results in more accessible and affordable care. Significant resources may be required to scale systems in health care and policymakers must take steps to remedy the uneven distribution of resources and access. There are varied applications of AI systems in health care such as research, health administration and operations, population health, practice delivery improvement, and direct clinical care. Payment and incentive policies must be in place to invest in building infrastructure, preparing personnel and training, as well as developing, validating, and maintaining AI systems with an eye toward ensuring value. While AI systems should help transition to value-based delivery models by providing essential population health tools and enhanced scalability and patient support, in the interim payment policies must incentivize a pathway for the voluntary adoption and integration of AI systems into clinical practice as well as other applications under existing payment models.

For example, when considering Medicaid and CHIP payment policy changes to improve access to AI-driven solutions in healthcare, CHI urges states to note that the American Medical Association's (AMA) Current Procedural Terminology (CPT) Editorial Panel has accepted a new Category 1 CPT® code for automated point-of-care retinal imaging (9925X) based on the recommendation of the American Academy of Ophthalmology. CHI submits that this development is a bellwether for the future of healthcare payments in light of increasing challenges in healthcare delivery and the benefits of AI and may serve as a model for Medicaid systems and CHIP seeking to responsibly enable AI tools to better serve their beneficiaries.

CHI further recommends that state Medicaid and CHIP policymakers evaluate whether their systems' administrative activities that utilize AI solutions using standardized and interoperable data formats may be eligible for enhanced federal payement match rates,[xx] and pursue such funding matches to enable more efficient and rapid adoption of AI.

Ethics: Policymakers should embrace the many of ethical norms emerging out of the medical and biomedical ethics community, and promote broader adherence of these norms by technologists, innovators, computer scientists, and those who use such systems. Healthcare AI will only succeed if protects patients and consumers. Policy frameworks should:

- Ensure that healthcare AI solutions align with all relevant ethical obligations, from design to development to use.
- Encourage the development of new ethical guidelines to address emerging issues with the use of AI in healthcare, as needed.
- Ensure consistency with international conventions on human rights.
- Ensure that AI for health is responsive to the sizable and growing systemic disparities in healthcare at all levels, and inclusive such that AI solutions beneficial to patients are developed across race, color, national origin, sex, age, disability, and other groupings.

- Reflect that AI for health tools may reveal extremely sensitive and private information about a patient or may reflect an ultimately inaccurate prediction of future risk (due to a mistake by the system or through targeted intervention in response to the prediction). States should ensure that laws protect and ensure that laws protect such information from being used to discriminate against patients.

Modernized Privacy and Security Frameworks: New uses and ways of analyzing healthcare data also raise new privacy questions and create new opportunities for more powerful and granular access controls for patients. Accordingly, any policy framework should address the topics of privacy, consent, and modern technological capabilities as a part of the policy development process. Privacy frameworks for healthcare AI should be scalable and assure that an individual's health information is properly protected, while also allowing the responsible and secure flow of health information to provide and promote high-quality healthcare and to protect the public's health and well-being. There are specific uses of data that require additional policy safeguards, i.e., genomic information. Given that one individual's DNA includes potentially identifying information about even distant relatives of that individual, a separate and more detailed approach may be necessary for genomic privacy. Further, enhanced protection from discrimination based on pre-existing conditions or genomic information may be needed for patients. Finally, with proper protections in place, policy frameworks should also promote data access, including open access to appropriate machine-readable public data, development of a culture of securely sharing data with external partners, and explicit communication of allowable use with periodic review of informed consent.

Collaboration and Interoperability: Policy frameworks should create a culture of cooperation, trust, and openness around health data that brings together policymakers, health AI technology developers and users, and the public. State policymakers should seek to leverage both the health datasets they collect, as well as SDOH, particularly with respect to ongoing public health crises (e.g., the COVID-19 public health crisis). We also recommend that states develop best practices to enable sharing of patient information within, and between, states in a safe and effective manner. Further, patients should be able to share and attain their own health data seamlessly, consistent with their expectations of data portability.

Workforce Issues and AI in Healthcare: All states face significant demands on the healthcare system and safety net programs due to an aging population and a wave of retirements among practicing care workers, with 'relatively fewer young people entering the workforce. Successful creation and deployment of AI-enabled technologies that help care providers meet the needs of all patients will be essential in addressing this projected shortage of care workers. Policymakers and stakeholders will need to work together to create the appropriate balance between human care and decision-making and augmented capabilities from AI-enabled technologies and tools.

Bias: Data biases exist when the AI's model or statistics are unrepresentative of a population, when data elements themselves are biased (e.g., physician-recorded levels of pain), or when labels reflect underlying bias. The bias inherent in all data as well as errors will remain one of the more pressing issues with AI systems that utilize machine learning techniques, and must be mitigated in all ways possible to address increasingly obvious disparities in healthcare. In developing and using healthcare AI solutions, these data provenance and bias issues must be addressed. Policy frameworks should (1) require the identification, disclosure, and mitigation of bias while encouraging access to databases and promoting inclusion and diversity, and (2) ensure that data bias does not cause harm to patients or consumers.

Lower income, minority, disabled, and other disadvantaged populations are often under-represented in data sets, yet represent significant parts of Medicaid and CHIP populations, which should be addressed when crafting AI solutions for any Medicaid and/or CHIP system. For example, AI-enabled clinical decision support tools used for the Medicaid population may need an underlying data set, improved through transparency measures, that accounts for the unique populations it is intended to serve.

Education: Policy frameworks should support expanding AI educational opportunities for the healthcare community and the patients they serve. Patients and consumers should be educated as to the use of AI in the care they are receiving and their rights and privacy options.

Academic/medical education should include curriculum that will advance healthcare providers' understanding of and ability to use health AI solutions. Ongoing continuing education should also advance understanding of the safe and effective use of AI in healthcare delivery.

V. Conclusion

CHI encourages all states to take meaningful steps to responsibly phase in new AI innovations into their health systems across contexts, consistent with the principles and recommendations above.

End Notes

[i] https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf

[ii] Oracle, What is Artificial Intelligence?, ORACLE SOLUTIONS, (last visited April 12, 2020), https://www.oracle.com/artificial-intelligence/what-is-artificial-intelligence.html.

[iii] McKinsey Global Institute, Artificial Intelligence: The Next Digital Frontier? (June 2017), available at https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx.

[iv] https://www.microsoft.com/en-us/ai/ai-for-accessibility.

[v] https://www.medicaid.gov/about-us/index.html

[vi] https://www.medicaid.gov/medicaid/index.html

[vii] https://www.medicaid.gov/chip/index.html

[viii] https://www.medicaid.gov/chip/index.html

[ix] https://hbr.org/2020/04/how-hospitals-are-using-ai-to-battle-covid-19.

[x] https://www.wired.com/story/opinion-ai-can-help-find-scientists-find-a-covid-19-vaccine/.

[xi] http://www.ihi.org/engage/initiatives/tripleaim/pages/default.aspx.

[xii] Thomas Bodenheimer, MD and Christine Sinsky, MDFrom Triple to Quadruple Aim: Care of the Patient Requires Care of the Provider, Ann Fam Med November/December 2014 vol. 12 no. 6 573-576.

[xiii] For example, the Centers for Disease Control and Prevention has noted inadequate reporting on racial disparities in coronavirus patients, which experts believe has hampered the public health response in communities of color. See https://appropriations.house.gov/events/hearings/covid-19-response-0.

[xiv] Defined as "an approach [that] focuses on interrelated conditions and factors that influence the health of populations over the life course, identifies systematic variations in their patterns of occurrence, and applies the resulting knowledge to develop and implement policies and actions to improve the health and well-being of those populations." Kindig, D. and Stoddart, G. What Is Population Health? American Journal of Public Health, 93, 380-383 (2003).

[xv] Nicole Lewis, Artificial Intelligence to play key role in population health, Medical Economics (2017) (available at http://www.medicaleconomics.com/medical-economics-blog/artificial-intelligence-play-key-role-population-health).

[xvi] https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/ForecastSummary.pdf.

[xvii] New Findings Confirm Predictions on Physician Shortage, ASSOCIATION OF AMERICAN MEDICAL COLLEGES, (April 23, 2019), https://www.aamc.org/news-insights/press-releases/new-findings-confirm-predictions-physician-shortage.

[xviii] Accenture, Artificial Intelligence: Healthcare's New Nervous System (2017), available at https://www.accenture.com/_acnmedia/PDF-49/Accenture-Health-Artificial-Intelligence.pdf#zoom=50.

[xix] Woolhandler et al, Costs of Health Care Administrationin the United States and Canada, N Engl J Med 2003; 349:768-75.

[xx] See https://www.macpac.gov/federal-match-rates-for-medicaid-administrative-activities/.

ConnectedHealth

Machine Learning and Medical Devices:
Connecting practice to policy *(and back again)*

By Sebastian Holst with Morgan Reed and Brian Scarpelli

Machine Learning and Medical Devices

Connecting practice to policy *(and back again)*

Contents

Introduction

Today, there are already many examples of artificial intelligence (AI) systems, powered by streams of data and advanced algorithms, improving healthcare by preventing hospitalizations, reducing complications, decreasing administrative burdens, and improving patient engagement. AI systems offer the promise to further accelerate and scale such results and provide impetus to the ongoing transition from our current disease-based system to one that is centered upon prevention and health maintenance. Nonetheless, AI in healthcare also brings with it a a variety of unique considerations for U.S. policymakers, particularly for medical device regulators.

Many organizations are taking steps to proactively address adoption and integration of AI into health care and how it should be approached by clinicians, technologists, patients and consumers, policymakers, and other stakeholders. Building on these important efforts, the Connected Health Initiative's (CHI) Health AI Task Force has taken the next step to address the role of AI in healthcare through the development of health AI policy principles.[1]

Generally, CHI believes that AI systems deployed in healthcare must advance the "quadruple aim" by improving population health; improving patient health outcomes and satisfaction; increasing value by lowering overall costs; and improving clinician and healthcare team well-being.

In order to succeed, Health AI systems must:

- Enhance access to health care.
- Empower patients and consumers to manage and optimize their health.
- Facilitate and strengthen the relationship and communication that individuals have with their health care team.
- Reduce administrative and cognitive burdens for patients and their health care team.

In providing its health AI policy principles with various key US federal policymakers, CHI's diverse AI Task Force has identified an opportunity to expand its contribution through a projection of its health AI policy principles onto a collection of good machine learning practices (GMLPs). Through a variety of public and collaborative initiatives designed to refine and build consensus around GMLPs, the objective is to provide a baseline that the Food and Drug Administration (FDA) an other governmental and non-governmental stakeholders can leverage in their their ongoing consideration of the topic. We intend for this document to serve as a next step in shaping health AI-related policy developments at the FDA, at the US federal level widely, and internationally.

CHI's AI Task Force welcomes collaboration with any interested stakeholder moving forward and appreciates consideration of this document.

[1] Connected Health Initiative *Policy Principles for Artificial Intelligence in Health*, https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf.

Effective governance is required to accelerate and amplify continued Machine Learning innovation

> ***The fastest cars need the best brakes.*** *To have the confidence required to drive at the highest speeds, a driver must trust their brakes – not just for emergencies, but for every scenario and under all conditions. And, without exception, the best brakes are engineered into the car; never added on as an afterthought*[1]*.*

Machine Learning[2] has advanced the quality and efficiency of medical devices and promises still greater innovations at an ever-quickening pace. Machine Learning's track record coupled with sky-high expectations for the future have also spawned a proportionate demand for – and investment in – effective governance; a means of assessing Machine Learning (ML) application suitability and performance, managing associated risks, and ensuring public safety and ethical use.

This document focuses on governance with respect two primary ML system categories: continuously learning systems (CLS) that are inherently capable of learning from real-world data and are able to update themselves automatically over time while in public use and "locks down" systems that have no ability to alter their configuration once testing and certification have been completed.

Governance strives to ensure appropriate levels of transparency, reliability, safety, security, and privacy.

Effective governance delivers on these objectives without compromising utility, efficiency, or innovation.

Effective ML governance is further required to instill confidence and trust in overall quality that, in turn, will lead to increased development velocity and ever-more ambitious innovation.

ML governance must be engineered into ML development practices and account for ML application behaviors

ML software behaves differently than traditional software in large part because it is developed differently.

[2] CHI supports the exemplary work of numerous organizations that are addressing healthcare AI, and seeks to harmonize and build upon these efforts including reuse, wherever possible, of accepted and recognized terminology and definitions. Unless defined inline, this paper will reuse the terminology and definitions included in in the December 2019-released Xavier University paper Building Explainability and Trust for AI in Healthcare. https://www.xavierhealth.org/news3/2020/1/8.

[3] This analogy has been borrowed with gratitude from the Open Compliance and Ethics Group, a non-profit think tank that promotes Principled Performance as the universal goal of every organization, team and individual.

Training Data shapes ML application behavior

Rather than explicitly define each logical sequence through source code as a traditional developer would, a ML developer transforms a generic predictive engine (an untrained machine) using a carefully curated training data set. In much the same way that a sculptor creates a mold around an original object, the ML developer creates a trained machine around a training data set. The training data set is constructed by the developer, but the training (computational analysis and resulting modifications to the untrained machine) are executed without developer intervention. The training data set has replaced source code at this stage of the development process and represents a wholly new development artifact.

How should training data sets be created, curated, and vetted?

Source code does not predict ML application behavior

There is no longer a one-to-one connection between application logic (behavior) and authored code. Depending on the training data set and the properties of the generic machine selected, the trained engine may have the ability to identify a broken bone in an X-ray, predict a heart attack, or dispense proper dosages of critical medication. Static analysis of peripheral source code or the training data set cannot predict the trained ML engine's behavior.

How can testing criteria be established if software behavior itself cannot be fully specified?

ML applications can continuously evolve

Unlike the compilation of source code into an executable program, machine training is not restricted to a single operation prior to an application's production release. If configured to do so, a trained machine that is in production (operational) can employ continuously learning systems (CLS) e.g. continue training using data consumed while in a production environment. This allows for the possibility that different copies of a single trained machine may each evolve independently from one another and from the initial trained machine.

How should new behaviors be evaluated in the field? When can this behavior even be safely deployed?

Effective governance of ML-enabled solutions begins with effective governance of ML software development and operations

The scale, complexity and distribution of ML applications has made governing each ML application instance recommendation, prediction, and action impossible.

What is possible – and practical – is to identify ML-specific risk factors stemming from the "paradigm-shifting" properties outlined above and evaluate how these have been proactively and transparently mitigated *within a broader software development lifecycle management context.*

It's not the "what", it's the "how."

The FDA Food Code ensures food safety and protection by focusing on broad areas of risk including the provisioning, preparation, and delivery of food.

It is not possible to evaluate each of the billions of food servings delivered every day. Governing the food supply chain and preparation "lifecycle" is the only practical means of effective governance.

How to get an A grade in ML software development

"FDA will assess the culture of quality and organizational excellence of a particular company and have reasonable assurance of the high quality of their software development, testing, and performance monitoring of their products[2]."

Broad Risk Categories	
Food	**Machine Learning**
Food from unsafe sources	Training data set deficits
Inadequate cooking	Machine training errors
Improper holding temperatures	Pipeline and distribution failures
Contaminated equipment	Operational vulnerabilities
Poor personal hygiene	Poor training and culture

Engineer effective ML governance into Medical Device software development lifecycles

There is an established practice of adapting vetted quality system management and software development lifecycle practices to support the unique priorities and requirements of the medical device industry.

The operative word here is "vetted." Due in large part to the three paradigm-shifting properties of ML technology outlined above, general ML software quality and development practices may be, in some circumstances, less mature than the development practices currently in place. The potential immaturity of some ML quality and risk management practices suggests that something more than "adapting" generally accepted practices will be necessary.

Given the accrued history and expertise of today's healthcare software developers – and SaMD developers in particular – this community has a material contribution to make in advancing – not merely adapting – mainstream development best practices.

[4] Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)

Part 1: Trace ML-specific properties through the software development lifecycle

The first task is to consider where traditional software development and quality management practices are most likely to require ML-specific accommodations prior to suggesting follow-on medical-device-specific adjustments.

The approach taken here is to trace ML-specific properties through the software development lifecycle. In much the same way that a contrast MRI employs a dye to highlight specific and difficult to detect conditions, this paper traces ML-properties across three interwoven software development axes with a special sensitivity to healthcare's overriding priorities, e.g. safety, transparency, and accuracy. The three development axes are:

1. Software manufacturing (the general principles of how whatever is developed is constructed, delivered, and maintained),
2. Software quality management (how suitability of purpose is defined and assessed for what is manufactured), and
3. Software security and risk management (frameworks and practices for identifying, assessing, and mitigating risks stemming from missed manufacturing or quality management requirements).

A Contrast MRI

A contrast MRI uses the injection of a contrast dye to better highlight certain conditions that might otherwise go undetected.

Part 2: Work-In-Progress Review: A Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device

In April of 2019, The FDA published an ambitious work that incorporated ML-centric principles into existing software development practices[5], Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback.

FDA U.S. 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

The stated goal was to advance a framework that would allow the FDA's regulatory oversight to embrace the iterative improvement power of machine learning for Software as Medical Device while assuring that patient safety is maintained.

Safety assurance is achieved through a multi-pronged approach that includes recommendations that ensure ongoing ML algorithm changes are:

- Implemented according to pre-specified performance objectives,
- Follow defined algorithm change protocols,
- Utilize a validation process that is committed to improving the performance, safety, and effectiveness of AI/ML software, and

[5] The authors acknowledge their debt to the International Medical Device Regulators Forum (IMDRF) for their work on SaMD (which, itself, relies upon prior IEC and ISO standards and frameworks) while recognizing the need for a "new, total product lifecycle (TPLC) regulatory approach that facilitates a rapid cycle of product improvement and allows these devices to continually improve while providing effective safeguards."

- Include real-world monitoring of performance.

These recommendations are rolled into an updated Total Product Lifecycle (TPLC) regulatory framework with the ultimate aim of promoting a mechanism for manufacturers to be "continually vigilant in maintaining the safety and effectiveness of their SaMD," supporting "both FDA and manufacturers in providing increased benefits to patients and providers."

As with The Food Code, the FDA would assess the culture of quality and organizational excellence of a particular company in order to establish "reasonable assurance" of the high quality of their software development, testing, and performance monitoring of their products.

The FDA, Machine Learning & SaMD

The FDA's has already begun the complex task of reimagining regulatory oversight to best embrace the power of machine learning while continuing to assure patient safety.

Given that general-purpose software development practices are themselves undergoing a material ML-driven evolution,

- Are there any underlying assumptions regarding quality and audit that merit closer review?
- What assurances can be built-in to ensure that those changes will be appropriately reflected in the central regulatory notions of "a culture of quality and excellence" and "reasonable assurance?"

Part 3: Beyond the Total Product Lifecycle[6]

Are there untapped approaches to embrace ML's most dynamic and opaque (but potentially powerful) properties? Are there longer-term opportunities to reimagine certification and pre-certification roles and workflows to further leverage AI/ML innovations?

Perhaps the most radical ML property from a regulatory perspective is the potential for algorithms to evolve after release and distribution. This capability is what is referred to as continuously learning systems.

Currently, this is only a theoretical concern as there is a blanket prohibition of this scenario across every existing and proposed TPLC regulatory framework.

Might there come a time when this prohibition will be perceived as imposing an undue constraint on innovation? Is there a scenario – perhaps in a robotics context – where allowing an initial set of SaMD instances to evolve wholly independently from one another will be identified as an absolute requirement? How would today's notions of manufacturing lifecycle and quality need to adapt?

Only "frozen algorithms" need apply (for now)

As with a graduating class of identically trained physicians whose skills mature independently over time, it is possible for an initial set of ML SaMD instances to evolve wholly independently from one another after distribution.

Might there come a time when the prohibition of real-time, continuous learning is perceived as an undue constraint on innovation?

[6] See Appendix C: Beyond the Total Product Lifecycle

Machine Learning is not the only transformative computing force. Cloud services, mobile 5G, and blockchain are among a growing list of revolutionary technological domains that are enabling entirely new ways of working, collaborating, and communicating.

Are there near-term organizational or technological opportunities that can help to prioritize near-term ML regulatory, governance and compliance requirements while also better positioning stakeholders across the healthcare and technology spectrum to capitalize on what may appear at first to be ML's most radical properties?

Tracing Machine Learning development properties through a general software development and DevOps lifecycle

Healthcare software governance combines policies and controls to:

- Ensure public safety
- Mitigate risks stemming from
  - Unintended consequences
  - Poor execution
  - Adversarial exploitation
- Encourage innovation in applications as well as the specialized development and testing tools required to produce those applications.

In what ways might ML development properties challenge foundational assumptions underlying traditional development lifecycle management practices?

Software Development Lifecycle Management

Software Development Lifecycle (SDLC) Management and DevOps tooling and practices normalize and automate software manufacturing processes while helping to ensure that safety, transparency, and privacy requirements are met.

In order for Machine Learning to complete its transition from paradigm-shifting innovation to a mainstream technology, SDLC management must also meet any additional requirements stemming from ML data-driven machine training development practices, e.g. Machine Learning Software Development Lifecycle Management (MLDLC).

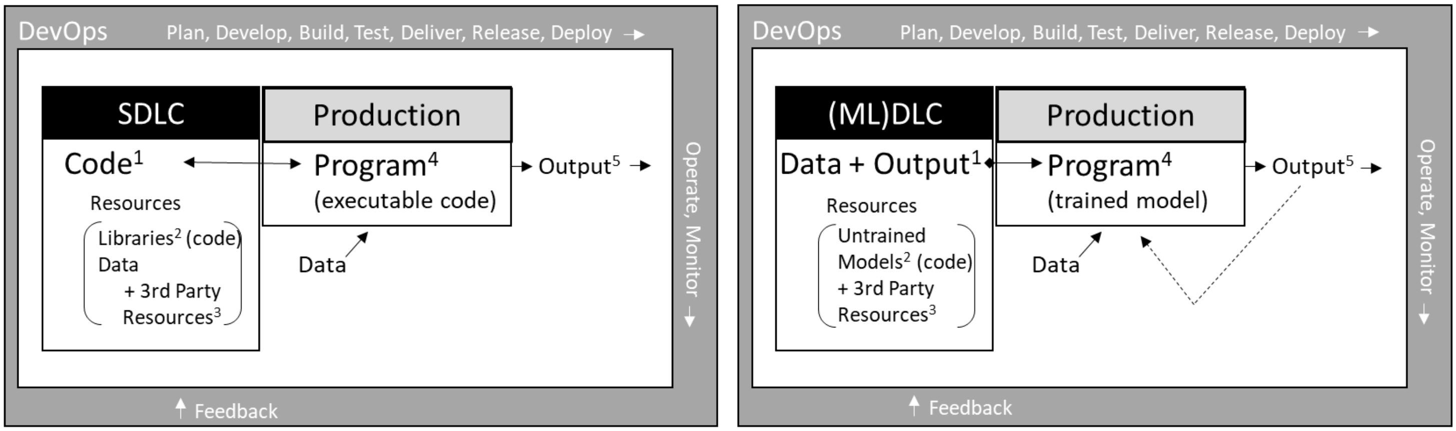


Figure 1: Traditional SDLC versus Machine Learning MLDLC wrapping in a DevOps iterative pipeline.

Figure 1 illustrates the elements of, and relationships between, a traditional Software Development Lifecycle and a Machine Learning Development Lifecycle operating within a well-formed DevOps pipeline.

The Figure 1 notes are described in the following table.

ML Key		Topic	Note
1 Training Data Set 2 Non-deterministic	**1**	Code vs Data + Output	Code sits at the center of a traditional SDLC and, consequently, is subject to rigorous quality, audit, and sourcing controls. Given that Data + Output supplants Code in a MLDLC, it follows that *an equivalent – but not identical – collection of controls are needed to ensure that effective quality, audit and sourcing remain in place.*
1 Training Data Set	**2**	Libraries vs Untrained Models	A traditional SDLC has built-in support for managing reusable code, typically in the form of libraries, to speed and simplify development, improve quality and auditability, and to help ensure consistency over time and across development teams. In much the same fashion, MLDLC will draw from a collection of reusable untrained models[7]. These models are code-based and are often organized as a traditional library, but given their heightened impact on development outcomes, *a corresponding increase in Untrained Model governance may also be justified.*
1 Training Data Set	**3**	3rd Party Resources	Today's applications increasingly rely upon 3rd party managed services, libraries, and software components. SDLC tools (Integrated Development Environments or IDE's) as well as software and service distribution channels have been extended to better support this rapidly evolving software supply chain. Supply chain risk management has also evolved to ensure appropriate visibility and accountability as the sourcing of code and services become increasingly distributed and diverse. *IDE's and IT security and risk management frameworks must evolve in-kind to keep pace with the consequences of including 3rd party Data + Output and/or Untrained Models into the modern software supply chain.*
1 Training Data Set	**4**	Production Programs	The traditional SDLC deliverable is an executable program. The MLDLC deliverable is a trained model. Due to ML statistical techniques, it is typically not possible – or nearly impossible – to trace exactly why a trained model behaves as it does. The absence of a decision tree in an ML program renders traditional SDLC code reviews, debugging, and general monitoring techniques obsolete. *ML programs may require compensating mechanisms to ensure comparable degrees of transparency, reliability and auditability.*
3 Incr. Learning	**5**	Output	Both traditional SDLC and DevOps best practices include a feedback loop that can be used to generate new requirements or improve existing features. This kind of continuous feedback fuels future program iterations and is subjected to the complete SDLC beginning with requirements through coding, test, etc. However, there are some branches of Machine Learning, specifically Continuously Learning where feedback is delivered directly into the current Production ML Program. These classes of Machine Learning bypass traditional SDLC inspection and approval steps and may result in unplanned and, potentially, unexpected behaviors. *Owners and regulators of sensitive and high-risk applications that must include human inspection may need to consider a blanket prohibition of these subcategories of Machine Learning until new norms about acceptable risk and transparency can be established. At a minimum, a greater understanding of the limitations and side-effects of deployed machine learning algorithms will be required by auditors and regulators.*

Table 1: MLDLC requirements stress traditional SDLC practices.

[7] ML programs also include "traditional reusable code" as well.

Machine Learning SDLC Requirement Summary

Tracing ML properties through high level SDLC stages suggested several potential new or modified requirements including:

1. The transition from code-driven to data-driven development will require corresponding practices and controls to meet quality, audit, and sourcing requirements.
2. Reusable Untrained Models are a special class of reusable code that, given their heightened impact on development outcomes, require a proportionate increase in governance.
3. Security and risk management must evolve in-step to keep pace with the implications of including 3rd party Data + Output and/or Untrained Models into the modern software supply chain.
4. Production ML programs may require novel monitoring and debugging mechanisms to ensure acceptable transparency, reliability, and auditability
5. Owners and regulators of sensitive and high-risk applications may need to consider blanket prohibitions of CLS Machine Learning models unless and until revised notions of transparency and predictability are established.
6. Integrated Development Environments (IDE's) and associated tooling will need to be extended to better scale and automate all phases of the new MLDLC.

Quality Management

While SDLC management measures and manages software manufacturing, distribution, and consumption, Software Quality is the field of study and practice that describes, measures, and manages the desirability (suitability) of the software itself.

Production Software Quality is, in large part, built upon Software Program Quality (the executable) that is, in turn, built upon the underlying Code Quality.

The shift to trained models away from code suggests a requirement to supplement existing code-centric quality practices and metrics.

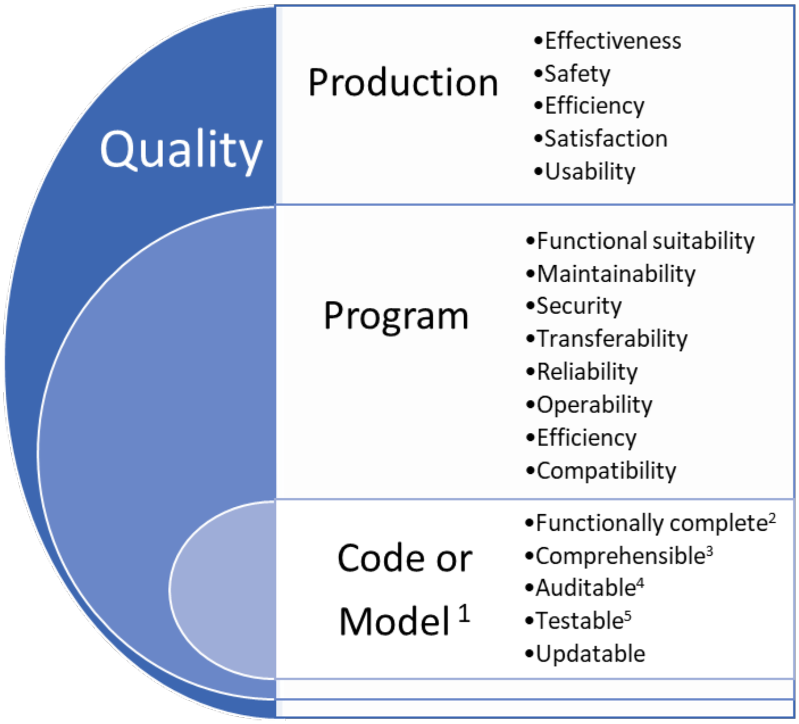


Figure 2: Quality is managed throughout the development lifecycle.

Figure 2 illustrates the elements of, and relationships between, common quality metrics divided into three segments: underlying code (or trained model), the resulting program, and the performance or suitability of that program.

Figure 2 notes are described in the following table:

ML Key		Quality Topic	Note
1 Training Data Set	**1**	Code vs Trained Model	Code sits at the center of a traditional Software Quality Practice with well-defined subcategories including functional completeness, comprehensibility, auditability, testability, and updatability. To preserve overall Quality, *ML development must develop equivalent – but not identical – methods of measuring and establishing acceptable quality metrics and tolerances.*
	Trained Model vs Code		
1 Training Data Set 3 Incr. Learning	**2**	Functionally Complete	Code can be statically analyzed, monitored for "coverage", and otherwise exercised to generate a mapping of input data and environmental states to expected outcomes. ML models are trained and tested through the processing of carefully curated data sets – there is no code that can be parsed and traced. Poorly formed datasets generate unexp ected and potentially unpredictable, behaviors and/or incorrect weighting of outcome predictions. Common examples of training data set gaps include: - Insufficient data volume - Lopsided data distribution across activities and outcomes - Missing activities and/or outcomes - Impossible activities or outcomes Poor data sets can result in the compromise multiple functional subcategories including: - Suitability: will the software behave appropriately for all users? - Accuracy: are functions implemented correctly? The models themselves may meet the highest quality standards, but the resulting trained model may fail to meet those standards. - Compliance: is the software in compliance with the necessary laws and guidelines? Transparency and predictability are required with virtually every regulatory and/or compliance obligation. *Development must have reliable means of detecting and, as needed, remediating gaps and other data set irregularities prior to ML model training.*
1 Training Data Set 3 Incr. Learning	**3**	Comprehensible	Every ML model includes intrinsic limitations. Understanding the stated purpose and objectives of a ML application and the hosting platform and implementation language will not be sufficient to assess the suitability of either training data or the selected ML models. In order to meaningfully "comprehend" the expected behavior of a trained *model, a reviewer must have specialized data science expertise and be knowledgeable in the strengths and limitations of the applied model(s) and the data staging/cleansing/sampling techniques.*
1 Training Data Set 3 Incr. Learning	**4**	Auditable	Tracing, reverse-engineering, and predicting how a model will behave given a specific set of inputs is difficult and, in practical terms, often impossible. This is especially true with extremely complex systems with many thousands of variables; the most common examples include image recognition, robotics, and natural language processing. *A consensus on acceptable alternatives to traditional event logging in code-based applications are needed to provide a comparable degree of assurance.* Untrained models are often provided by open source communities or platform providers. *A common format for sourcing the precise model and version with a record*

			of know Quality issues would help to predict Quality issues that may arise in the final trained model.
1 Training Data Set 1 Training Data Set 3 Incr. Learning	**5**	Testable	Exception detection, defect definition, and related KPI's (including testing cost) must be established to effectively model the severity and cost of ML application defects specifically related to under-performance. Output measurement must also be standardized, utilizing what developers measure for their own data models including terminology and their own interpretation of medical information*. This industry-specific formulation results in a harmonization of terminology across regulators and stakeholders that will improve quality management.*

Table 2: Trained Models drive expansion of code-centric Software Quality practices.

ML Software Quality Summary

Tracing ML properties across basic Quality System segments suggested several additional new or modified requirements including:

1. ML Software must meet the same quality standards as code-based software. As such, there must be equivalent methods of measuring and establishing acceptable ML-centric quality metrics and tolerances to offset inapplicable code-centric controls.
2. ML-centric controls must cover both the special data sets used for training and testing ML models as well as the trained ML models themselves.
3. Reviewers, testers, and auditors will require additional specialized data science expertise including a working knowledge of the strengths and limitations of deployed model(s), the implications of their parameters as well as any data staging/cleansing/sampling techniques that are applied.
4. The sourcing of untrained models is a potential supply chain gap – in much the same way that a revised compiler can introduce quality issues in established source code. A common format for sourcing a precise model and version with a record of known quality issues would likely help to predict Quality issues that may arise in a final trained model.
5. Quality Systems must also incorporate updated and harmonized health care specific terminology, data collection, and measurement practices to ensure the availability of relevant baseline healthcare quality metrics and standards.
6. The establishment of exception detection, defect definition, and related KPI's (including testing cost estimation) are needed to effectively model the severity and cost of ML application defects specifically related to ML under-performance.

Software Security and Risk Management

Effective risk and security management begins with identifying and prioritizing material threats and works to establish effective controls that reduce risk to acceptable levels. For application risk and security management, recommended practices typically include:

- Detailed Abuse Cases[8] that are used to
  - Develop a business/technical specific Threat Model[9] that in turn is used to assess risks stemming from
    - Each application's Attack Surface[10], e.g. the application's entry and exit points.

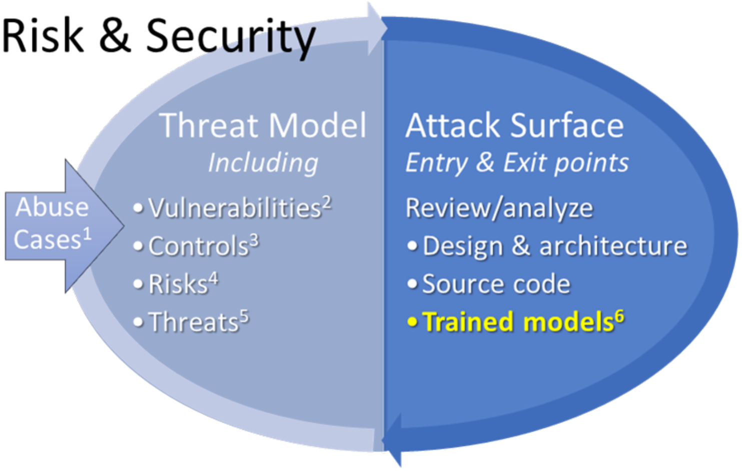


Figure 3: Risk and Security Modeling

These interrelated components evolve with production usage and feedback generating additional Abuse Cases that in turn update the Threat Model resulting in further refinements to the application's Attack Surface and underlying controls.

Software Security and Risk Management practices must also expand to meet new requirements stemming from Machine Learning development practices, technology, and use cases. Figure 3 notes are described in the following table.

ML Key

	Risk & Security	Note
1	Abuse Cases	*The current paucity of established ML Abuse Cases is likely to lead to an incomplete view of potential threats and undermine threat modeling activities and the subsequent control priorities that follow.*
2	Vulnerabilities	ML systems novel use of training data to create production behaviors have spawned an equally novel set of novel vulnerabilities including: • Data poisoning (injecting training data designed to cause errors) • Adversarial input (data crafted to be misclassified by targeted models) • Exploitation of errors in autonomous system goals *The set of known ML-specific vulnerabilities is almost certainly incomplete as are the range of potential exploits.*
3	Controls	*There is a further deficit in established Preventative and Detective Controls to mitigate the risks stemming from ML-inspired vulnerability attacks.*
4	Risks	Effective risk assessments are dependent upon accurate probability estimates. Risk calculations typically combine:

[8] OWASP Abuse Case Cheat Sheet https://github.com/OWASP/CheatSheetSeries/blob/master/cheatsheets/Abuse_Case_Cheat_Sheet.md

[9] OWASP Application Threat Modeling https://www.owasp.org/index.php/Application_Threat_Modeling#1._What_are_we_building.3F

[10] OWASP Attack Surface Cheat Sheet https://github.com/OWASP/CheatSheetSeries/blob/master/cheatsheets/Attack_Surface_Analysis_Cheat_Sheet.md

		• The probability of an incident occurring (an exploit of a vulnerability) • The probability of that incident causing harm and • The degree of harm that comes with each occurrence *The rapidly evolving use of ML across industries and use cases significantly complicate ML risk assessment calculations making risk mitigation investment decisions more difficult to calibrate.*
5	Threats	In addition to the exploitation of unique ML vulnerabilities, the weaponization of ML in the hands of bad actors must also be considered. Examples include: • Automation of social-engineering attacks and the dissemination of political misinformation leveraging improved profiling, messaging and deep fake image and audio generation. • Anonymization and scaling of physical assaults using autonomous drones and other vehicles • Highly efficient and distributed cyber-attacks leveraging specialized ML models. • Expansion of potential attackers as democratization of all of the above removes human domain expertise as a requirement. *ML expands the variety of potential threats, improves the efficiency of existing threats, and expands the number of potential attackers.*
6	Trained models	*ML training and test data sets represent additional attack surface opportunities to be included in current Attack Surface mapping practices.*

Table 3: Machine Learning impact on established Application Risk and Security practices

Machine Learning Security and Risk Management Summary

Tracing ML properties through security and risk management categories highlight some measure pf risk from all three ML property categories listed above.

1. The short history of successful ML exploits constrains Threat Modeling practices.
2. The inventory of ML-specific vulnerabilities is incomplete as are the understanding of potential exploits.
3. There is a further deficit in established Preventative and Detective Controls to mitigate the risks stemming from ML-inspired vulnerability attacks.
4. The rapidly evolving use of ML across industries and use cases significantly complicate ML risk assessment calculations making risk mitigation investment decisions more difficult to calibrate.
5. ML training and test data sets represent additional attack surface opportunities to be included in current Attack Surface mapping practices.
6. ML has a multiplicative effect on Risk and Security management by expanding the variety of potential threats, improving the efficiency of existing threat tactics, and expanding the number of potential attackers

Work-In-Progress Review: A Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device

A Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback was published with the stated goal of advancing a framework to allow the FDA's regulatory oversight to embrace the iterative improvement power of machine learning for Software as Medical Device while assuring that patient safety is maintained.

The proposed Total Product Lifecycle (TPLC) regulatory framework is designed to ensure ongoing ML algorithm changes are:

- Implemented according to pre-specified performance objectives,
- Follow defined algorithm change protocols,
- Utilize a validation process that is committed to improving the performance, safety, and effectiveness of AI/ML software, and
- Include real-world monitoring of performance.

In order to manage the scale and scope of this ambitious effort and to avoid the necessity of auditing every development milestone of every software component, the FDA proposes assessing the culture of quality and organizational excellence of a particular company in order to establish "reasonable assurance" of the high quality of their software development, testing, and performance monitoring of their products.

As outlined in the prior section, much of the underlying general-purpose software development standards, frameworks, and practices[11] are themselves actively undergoing their own ML-driven evolution. This section drills into the updated Total Product Lifecycle Regulatory approach and the associated "Culture of Quality and Organizational Excellence" to identify:

- Underlying assumptions regarding Software Development Lifecycle Management, Quality or Risk that may merit closer review, and
- Mechanisms to ensure evolving assumptions are appropriately reflected in the central notions of "a culture of quality and excellence" and "reasonable assurance."

In order to "balance the benefits and risks, and provide access to safe and effective AI/ML-based SaMD," the revised TPLC seeks to establish clear expectations on quality systems and good ML practices (GMLP) as outlined in the following illustration.

[11] See Appendix A: Supporting organizations and underlying standards and frameworks.

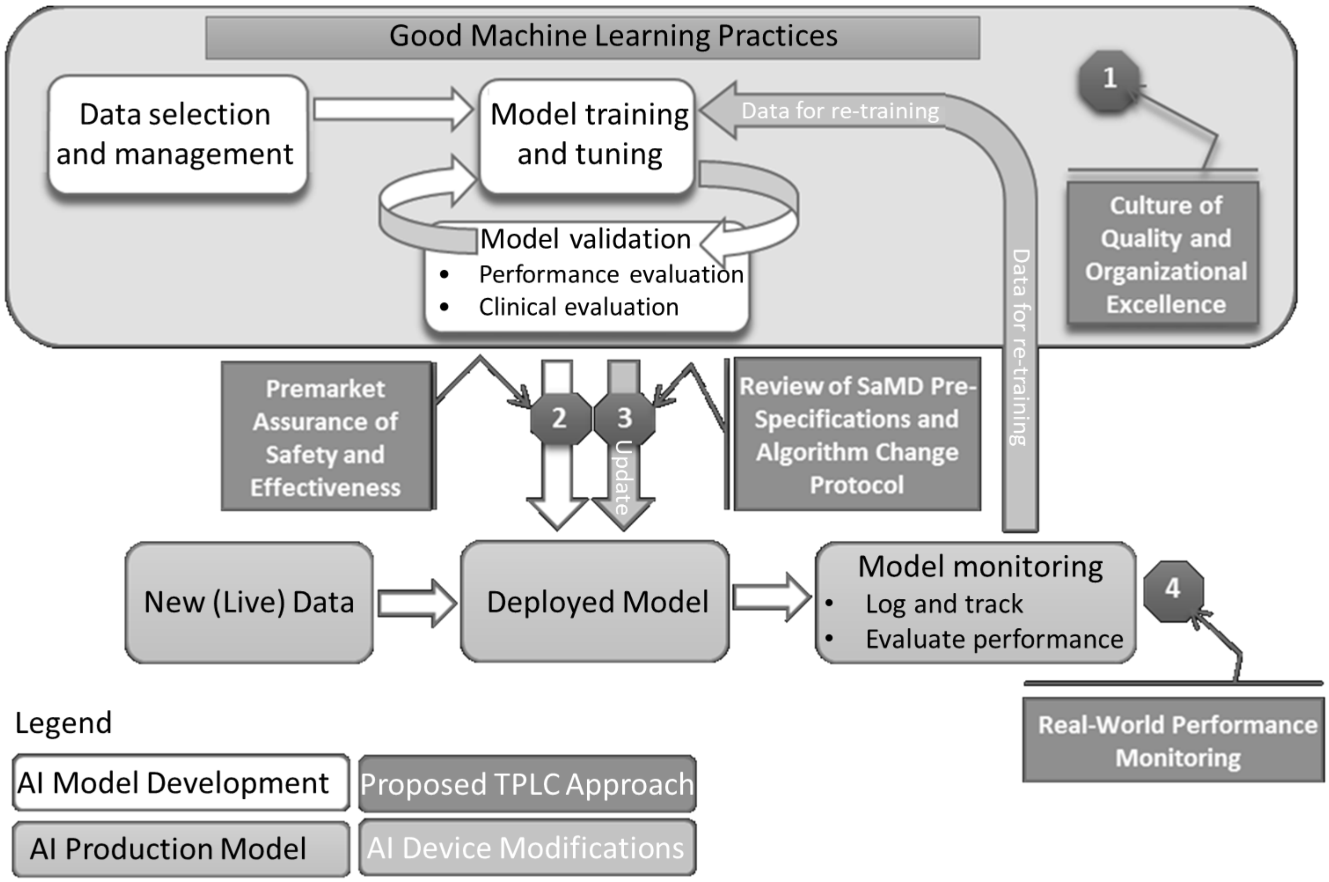


Figure 4: Overlay of FDA's TPLC approach on an ML workflow

Figure 4 notes are described in the following table.

ML Key	Note	TPLC-specific guidance is subject to underlying ML Development and Operations dependencies
1 Training 1 Training Data Set	1	A Culture of Quality and Organizational Excellence has historically relied upon IEC 62304 for establishing required "lifecycle support processes." *IEC 62304 is currently code-centric in its audit, test and monitoring assumptions.*
1 Training Data Set 3 Incr. Learning	2	Premarket Assurance of Safety and Effectiveness identifies ML-data-centric gaps *but specific patterns and practices have not (yet) been addressed.*
3 Incr. Learning	3	Review of SaMD Pre-Specifications and Algorithm Change Protocol works to constrain many of the dynamic, continuous adaptation capabilities of some ML algorithms in order to mitigate unexpected results in the field. *Without some breakthroughs in transparency and monitoring, many of the most dynamic learning algorithms will most likely be entirely prohibited for use inside SaMDs.*
1 Training 1 Training Data Set 3 Incr. Learning	4	Real-World Performance Monitoring is an essential to ensuring transparency, effectivity and actual usage patterns. *Special care must be taken to correctly interpret results as a measure of ML model performance and differences between SaMD model releases.*

Table 4: ML development considerations within the FDA's proposed Total Product Lifecycle Regulatory approach

GMLP Summary

Evaluating GMLP in the context of the ongoing evolution of ML-centered development quality, SDLC, and risk management, the following issues may merit deeper investigation:

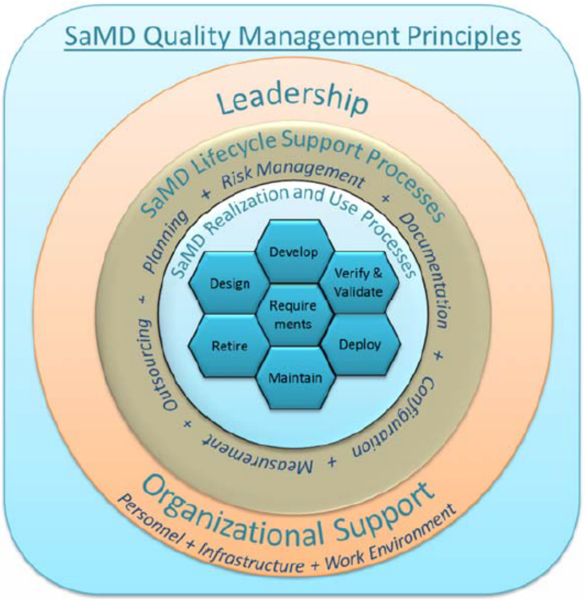
1. Heavy reliance on standards that have historically been defined by methodical and deliberate revision policies may not be able to keep pace with rapidly changing development practices and exacerbate rather than mitigate quality risk stemming from ML's data-driven versus code-driven properties.
2. Without a sufficient body of verified ML development patterns have been documented, it may be difficult to establish a durable definition of "reasonable" and "effective."
3. The long-standing requirement that all copies of a given device or software instance can only by updated but cannot independently evolve prohibits a subset of dynamic and continuously learning applications.
4. Incident management and platform monitoring systems will likely need to expand incident categories and severity ratings to account for unique classes of exceptions unique to ML services.

The Culture of Quality and Organizational Excellence

The Culture of Quality and Organizational Excellence is itself comprised of three management principles:

1. Leadership that sets the organizational tone,
2. Lifecycle Support Processes that wrap and operationalize the actual development, and at its core,
3. Deployment, and maintenance activities associated with actual SaMD development.

As noted in Table 4, note 1 above, software lifecycle standards, such as IEC 62304, are code centric and will likely need to be extended or adapted to the unique lifecycle requirements associated with training ML algorithmic models.



FDA SaMD QMS Principles

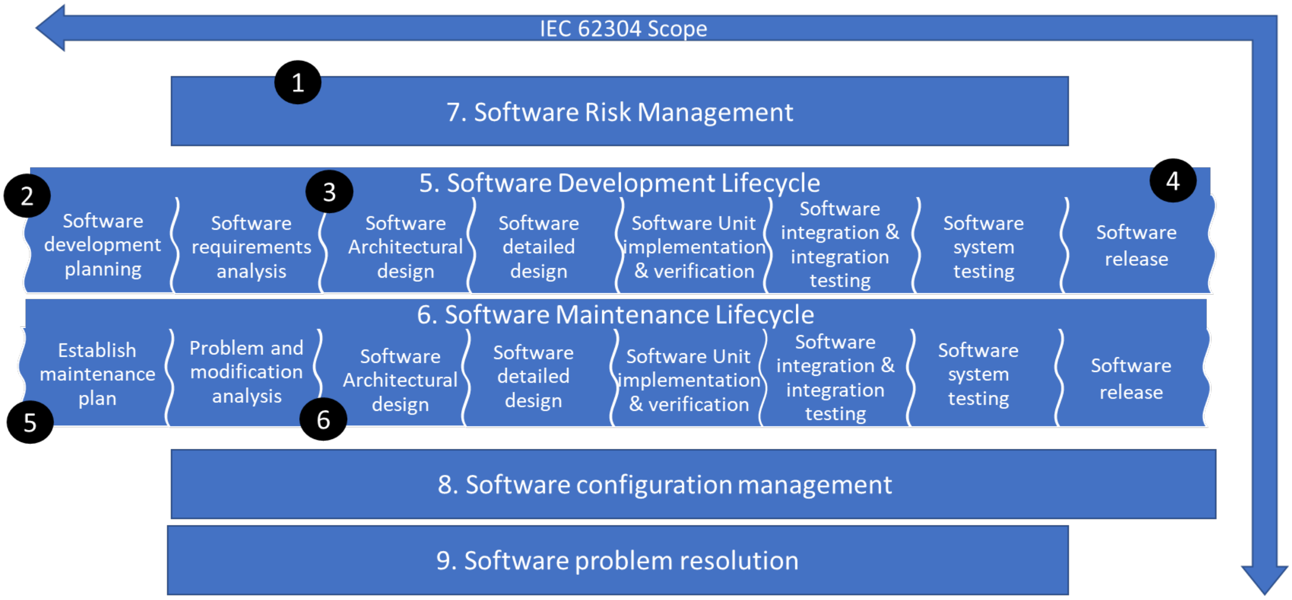

IEC 62304 Scope
1
7. Software Risk Management
2
3
4
5. Software Development Lifecycle
Software development planning
Software requirements analysis
Software Architectural design
Software detailed design
Software Unit implementation & verification
Software integration & integration testing
Software system testing
Software release
6. Software Maintenance Lifecycle
Establish maintenance plan
Problem and modification analysis
Software Architectural design
Software detailed design
Software Unit implementation & verification
Software integration & integration testing
Software system testing
Software release
5
6
8. Software configuration management
9. Software problem resolution

Figure 5: ML development impact on IEC 62304 development lifecycle processes.

Figure 5 notes are described in the following table.

Note	ML development impact on IEC 62394 development lifecycle processes.
1	Software Risk Management: *How will risks associated with training data sets be mitigated?*
2	Software development planning: *How will Software Of Unknown Providence (SOUP) be extended to accommodate 3rd party algorithms and external training data?*
3	Software requirements analysis: *How will issues relating to bias and transparency be incorporated?*
4	Software release: *Given the requirements above, how can FDA Premarket Safety Assurance requirements be effectively be met?*
5	Maintenance plan: *Defining, measuring, and documenting the degree of change within an SaMD will require significant coordination and consensus.*
6	Problem and modification analysis: *Documenting root causes and effectivity of modifications stemming from data set deficiencies will require new (or enhanced) concepts, tooling and terminology.*

Table 5: ML development considerations within IEC 62304: Medical device software lifecycle processes.

Culture of Quality and Organizational Excellence

Evaluating working definition of the Culture of Quality and Organizational Excellence in the context of the ongoing evolution of ML-centered development quality, SDLC, and risk management, the following issues may merit deeper investigation:

1. To satisfy an external auditor/examiner, Organizations will need to be able to tap into a sufficiently large body of recognized ML controls able to substantially meet their requirements.
2. Suppliers of third party and embedded software, also referred to as Software Of Unknown Providence (SOUP), must be able to satisfy corresponding requirements for transparency, safety, security, and privacy.
3. Individuals will need the ability to know if/how their data may be used to develop and/or train machines or algorithms. The opportunity to participate in data collection for these purposes must be on an opt-in basis.[12] [13]
4. A consensus must be reached on the definition and measurement of a wholly new quality criteria related to behavior, e.g. bias and human-readable decision-making transparency.
5. New (or enhanced) concepts, tooling and terminology will likely be required across a broad spectrum of operations management capabilities to properly capture the impact of dataset deficiencies including:
   - Chance control documentation including risks assessment,
   - Root cause analysis, and
   - Modification effectiveness.

[12] Connected Health Initiative *Policy Principles for Artificial Intelligence in Health*, https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf.

[13] American Medical Association's privacy principles https://www.ama-assn.org/system/files/2020-05/privacy-principles.pdf.

Initial observations

There is wide agreement that existing regulations need revision to accommodate the unique (and potentially disruptive) properties of Machine Learning technologies and development processes.

> *100% of Proposed Regulatory Framework responses endorsed the requirement to update existing medical device regulatory obligations to accommodate Machine Learning*[14]*.*

The FDA, responding to this need, has proposed a regulatory framework to manage what is likely to be one of the most challenging aspects of regulating ML-driven "Software as Medical Devices," modifications that may, or may not, require a review and recertification – a potentially time-consuming and expensive process.

> *One of the distinguishing properties of the Machine Learning approach is the capacity for programs to alter behavior over time without requiring additional coding or software updates. This kind of unsupervised learning challenges conventional development, quality, and risk practices and policies.*

The FDA proposal built off existing regulations, frameworks, and definitions, extended some where needed, and added wholly new constructs when it was determined to be unavoidable.

> *Initial feedback to the proposed framework reinforced the importance of leveraging existing standards and framework – perhaps to an even greater extent than the initial proposal envisioned.*

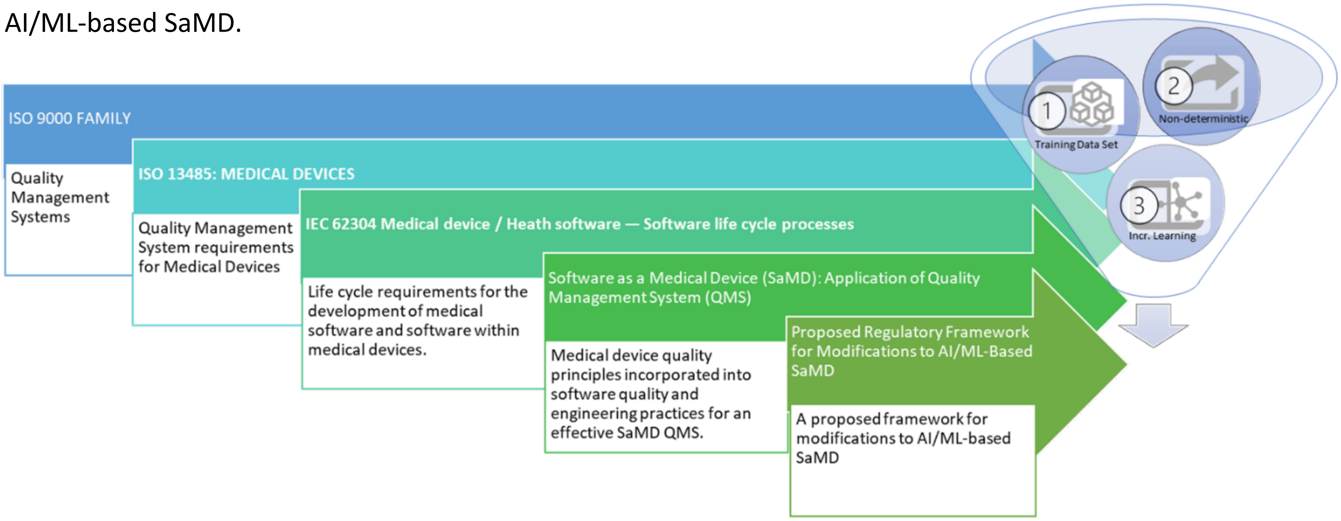
There is significantly more work that needs to be done refining and harmonizing definitions, completing core processes and performance metrics, as well as educating the vast community of stakeholders.

> *Tracing ML-specific development and technical properties from Innovator practices through relevant tooling, development frameworks, and standards promises to ultimately shorten and simplify the work required to effectively and efficiently "protecting the public health by ensuring Software as Medical Device safety, efficacy, and security."*
>
> *This can be most effectively accomplished through a sustained collaboration with, and communication across, the stakeholder ecosystem (innovators, platform providers, supranational standards bodies, government regulators, etc.).*

[14] See Appendix B: Respondent Submission Analysis

Appendix A: Supporting organizations and underlying standards and frameworks

There is an established practice of adapting vetted quality system management and software development lifecycle practices to support the unique priorities and requirements of the medical device industry. The following list includes frameworks and documents, as well as the associated governing organizations, that provide underlying support for the FDA's Proposed Framework for Modifications to AI/ML-based SaMD.


ISO 9000 FAMILY
Quality Management Systems
ISO 13485: MEDICAL DEVICES
Quality Management System requirements for Medical Devices
IEC 62304 Medical device / Heath software — Software life cycle processes
Life cycle requirements for the development of medical software and software within medical devices.
Software as a Medical Device (SaMD): Application of Quality Management System (QMS)
Medical device quality principles incorporated into software quality and engineering practices for an effective SaMD QMS.
Proposed Regulatory Framework for Modifications to AI/ML-Based SaMD
A proposed framework for modifications to AI/ML-based SaMD
1
Training Data Set
2
Non-deterministic
3
Incr. Learning

International Electrotechnical Commission (IEC)

The IEC prepares and publishes International Standards for all electrical, electronic and related technologies.

IEC 62304:2006/AMD 1:2015 Medical device software life cycle processes is a standard which specifies life cycle requirements for the development of medical software and software within medical devices.

International Organization for Standardization (ISO)

ISO is an independent, non-governmental international organization with a membership of 164 national standards bodies. ISO – in conjunction with the IEC – has identified the need to develop standards for AI that "can benefit all societies." Established in 2017, this is the charter of the ISO/IEC Joint Technology Committee (JTC) 1 / SubCommittee (SC) 42 for artificial intelligence (SC 42).

SC 42's scope includes basic terminology and definitions, risk management, bias and trustworthiness in AI systems, robustness of neural networks, machine-learning systems and an overview of ethical and societal concerns. SC 42 has already published three Big Data standards with 13 projects currently under development. Five of these are highlighted below.

ISO/IEC JTC 1/SC 42: Artificial Intelligence

AI/ML ISO standards under development from ISO/IEC JTC 1/SC 42 include:	
ISO/IEC 23053	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)
ISO/IEC 24027	Bias in AI systems and AI aided decision making
ISO/IEC 38507	Governance implications of the use of artificial intelligence by organizations
ISO/IEC 23894	Artificial Intelligence — Risk Management
ISO/IEC TR 24368	Artificial Intelligence (AI) — Overview of ethical and societal concerns

International Medical Device Regulators Forum (IMDRF)

The IMDRF is a voluntary group of medical device regulators from around the world who have come together to form the Global Harmonization Task Force on Medical Devices (GHTF) whose mission is to "accelerate international medical device regulatory harmonization and convergence." Their relevant works to date are highlighted here.

IMDRF publications include:	
IMDRF/SaMD WG/N10	SaMD: Key Definitions
IMDRF/SaMD WG/N12	SaMD: Possible Framework for Risk Categorization & Corresponding Considerations
IMDRF/SaMD WG/N23	SaMD: Application of Quality Management System
IMDRF/SaMD WG/N41	SaMD: Clinical Evaluation

US Food and Drug Administration (FDA)

The FDA is responsible for protecting the public health by ensuring the safety, efficacy, and security of drugs, biological products, *and medical devices*. In addition to the Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback, the FDA is also active in contributing to, endorsing, and re-publishing many of the IMDRF publications listed above. At this time, the FDA has not made ML-specific modifications to Medical Device regulatory obligations (see 21 CFR Parts 803 through 861).

Appendix B: Respondent Submission Analysis

Proposal Questions and Feedback

While there were no constraints placed on the kinds of feedback or questions that could be submitted, the FDA included questions that covered the most important (or perhaps controversial) elements of the proposed TPLC framework.

Questions included in Proposed Regulatory Framework were divided into subtopics.

- How complete is the classification of AI/ML SaMD modifications and will they be effective and helpful?
- Is the GMLP complete? How can the FDA help manufactures incorporate new requirements into their existing QMS systems and practices?
- All feedback to the definitions and implementation details surrounding SPS and ACP. These are entirely new elements to the proposed certification process.
- How can the process of premarket review (review prior to an initial SaMD launch) be better defined and managed?
- How can "real-world" data be captured, analyzed, secured, and weighted throughout this entire process?
- What should the ACP include and how can it be consistently and effectively assessed across manufacturers and SaMDs?

These questions bring to the fore just how potentially disruptive Machine Learning may be in the short-term – and why it is in everyone's interest to shorten the ML transition into the mainstream.

That being the case, why did 64% if respondents fail to answer even one of the FDA's questions?

64% of the public responses did not directly reference a single question included in the Framework Proposal.

Questions included in Proposed Regulatory Framework

The types of AI/ML-SaMD modifications (Key: **AI/ML SaMD**)

1. Do these categories of AI/ML-SaMD modifications align with the modifications that would typically be encountered in software development that could require premarket submission?
2. What additional categories, if any, of AI/ML-SaMD modifications should be considered in this proposed approach?
3. Would the proposed framework for addressing modifications and modification types assist the development AI/ML software?

Good Machine Learning Practices (**Key: GMLP**)

1. What additional considerations exist for GMLP?
2. How can FDA support development of GMLP?
3. How do manufacturers and software developers incorporate GMLP in their organization?

SPS and ACP **(Key SPS/ACT)**

1. What are the appropriate elements for the SPS?
2. What are the appropriate elements for the ACP to support the SPS?
3. What potential formats do you suggest for appropriately describing a SPS and an ACP in the premarket review submission or application?

Premarket review **(Key: PreMarket)**

1. How should FDA handle changes outside of the "agreed upon SPS and ACP"?
2. What additional mechanisms could achieve a "focused review" of an SPS and ACP?
3. What content should be included in a "focused review"?

The transparency and real-world performance monitoring (**Key: Transp & Monitoring**)

1. In what ways can a manufacturer demonstrate transparency about AI/ML-SaMD algorithm updates, performance improvements, or labeling changes, to name a few?
2. What role can real-world evidence play in supporting transparency for AI/ML-SaMD?
3. What additional mechanisms exist for real-world performance monitoring of AI/ML-SaMD?
4. What additional mechanisms might be needed for real-world performance monitoring of AI/ML-SaMD?

ACP Scope: **(Key: ACP)**

1. Are there additional components for inclusion in the ACP that should be specified?
2. What additional level of detail would you add for the described components of an ACP?

The following analysis is based upon the public responses to The Proposed Regulatory Framework for Modifications to AI/ML-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback.

Looking at the respondents' own questions and/or their interest (and/or lack of interest) in the FDA's questions offers insight into how stakeholders outside of the FDA perceive these issues and which of these may be perceived as more (or less) important or controversial.

Respondent industries and corresponding stakeholder community roles

Respondent submissions are available for review on the FDA website[15]. Figure B1 maps the self-identified Industry Categories of 127 respondents to generic Stakeholder Community roles[16].

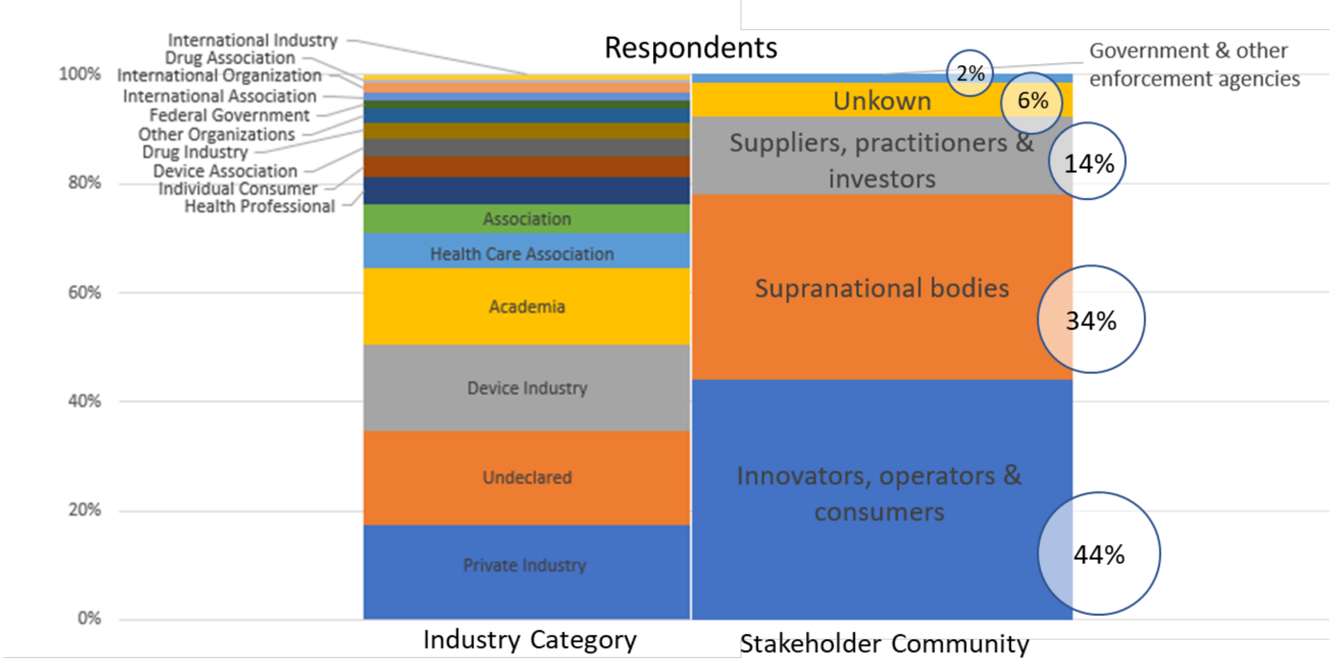


Figure B1: Respondent Industry Categories and Stakeholder Roles

Perhaps it is not surprising to learn that the primary stakeholders have the loudest voice (at least by sheer volume), but, given the importance of vendor-neutral, independent "Supranational bodies" in shaping regulations, should they?

Respondent priorities

The questions embedded inside the FDA's regulatory framework proposal are calibrated to address the FDA's priorities, but are those priorities and their relative weighting shared? Figure B2 illustrates the percentage of responses that included specific topics. These topics are grouped into "framework-specific" (that are unique to the proposed regulatory framework) and "mainstream activities" (that are general issues already described relating to the mainstreaming of any disruptive technology).

[15] https://www.regulations.gov/document?D=FDA-2019-N-1185-0001

[16] When included, the respondent's organization was also used to map into the Stakeholder Community role.

64% of respondents did not answer any of the 18 questions included in the proposal. Closer inspection of respondents' comments suggests a difference in emphasis and, perhaps, priority.

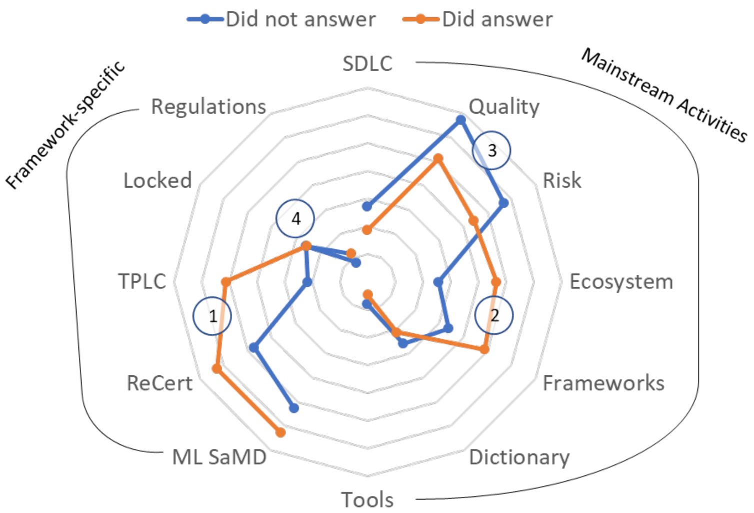


Figure B2: Topic interest of respondents

Respondents that did answer FDA-specific questions:

1. Were much more likely to comment on the ML SaMD modification categories, the recertification criteria and process, and the description of the TPLC.
2. Consistently raised issues across the mainstream activities of Quality, Risk, Ecosystem (collaboration across roles) and Frameworks (reconciliation with other frameworks).
3. Respondents that did not answer the FDA-specific questions were significantly more likely to focus on software Quality and Risk issues.
4. Regardless of whether the FDA-specific questions were addressed, there was a general concern around the definition and treatment of "Locked" models.

Respondent priorities by topic

Does a respondent's stakeholder role as innovator or standards body (versus regulatory agency or consumer) also influence their priorities? If yes, should the dominance of one stakeholder role over all others be factored-in or weighted when considering responses?

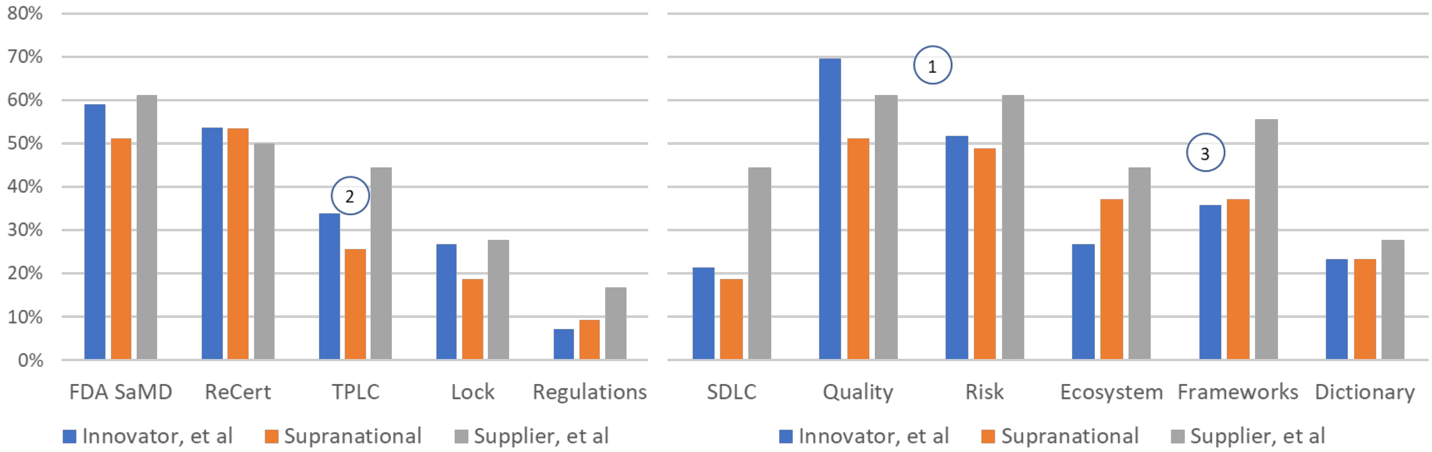


Figure B3: percentage of responses across topics by Ecosystem Stakeholder role.

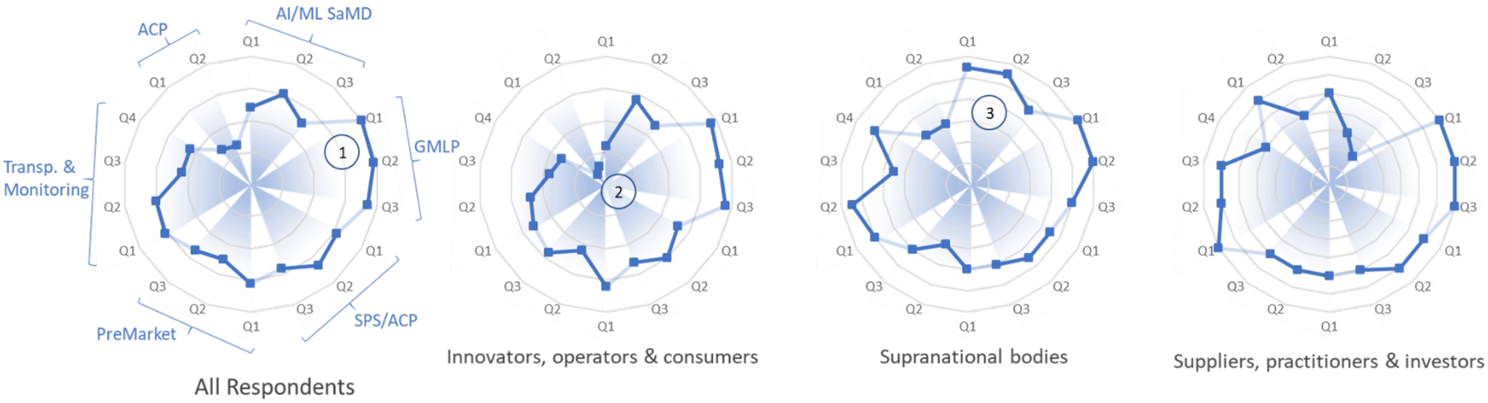
Figure B3 maps the percentage of topics included in responses by Stakeholder role (only three roles had enough responses to be statistically meaningful).

1. Quality, Risk, FDA SaMD modifications and recertification processes received the greatest attention.
2. Generally, Innovators, consumers, practitioners and suppliers responded more consistently with one another as compared to Supranational organization responses.

3. Taken as a group, comments relating to Ecosystem (cross roll collaboration), Frameworks (cross framework reconciliation), and Dictionary (defining common terms and definitions across domains) were a strong, consistent area of concern.

FDA-specific question response

While only 36% of respondents addressed the embedded 18 questions directly, those responses were extensive and, obviously, important to assess.



All Respondents

Innovators, operators & consumers

Supranational bodies

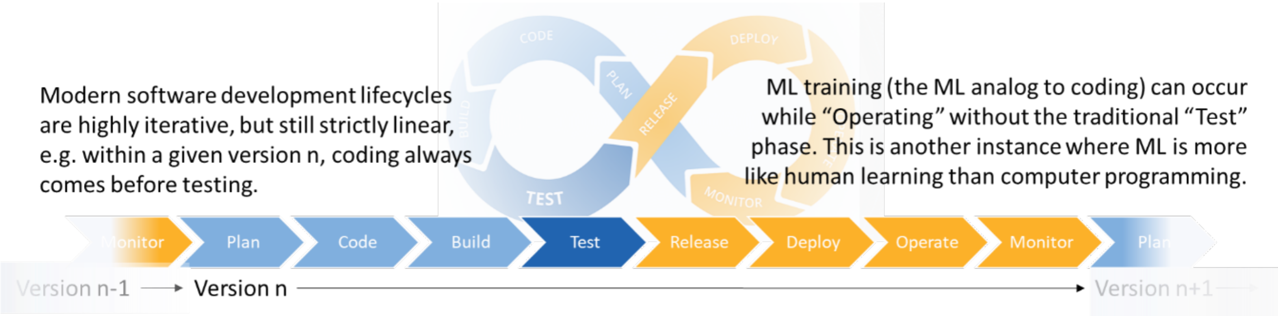
Suppliers, practitioners & investors

Figure B4: Count of responses that included commentary for each FDA-embedded question. The questions are segmented by topic. All Respondents are shown alongside the three highest reporting Ecosystem Stakeholder roles.

1. Respondents gave the greatest amount of attention to the questions relating to Good Machine Learning Practices.
2. Relative to the other subtopics, Algorithm Change Protocol received substantially less attention from Innovators, et al than any other subtopic. This gap was not evident in either of the other two Stakeholder roles.
3. The high innovator response volume depressed the relative importance of the ACP subtopic. Given the close relationship between Supranational Organizations and Government Regulators already discussed and the consensus around the importance of framework and regulatory consistency, should the (apparent) lack of interest from Innovators be discounted?

Appendix C: Beyond the Total Product Lifecycle

Software development lifecycle management, like virtually all modern Product Lifecycle Management, is a highly iterative process, but within any given version, the lifecycle stages are executed in a strictly linear sequence. As an example, within a given version n, coding, building, and testing must always precede deployment and production operation.


Modern software development lifecycles are highly iterative, but still strictly linear, e.g. within a given version n, coding always comes before testing.

ML training (the ML analog to coding) can occur while "Operating" without the traditional "Test" phase. This is another instance where ML is more like human learning than computer programming.

When configured to do so, continuously learning algorithms can breach the strict sequencing imposed by development lifecycle methodology. Not surprisingly, the FDA's proposed AI/ML TPLC includes a prohibition of this kind of evolutionary behavior in real-time and in production. This is a sound policy as there is no precedent to contradict this position to be found in the underlying standards and frameworks.

Yet, while there is no *underlying* precedent, might there be a precedent to be found in an *adjacent health care domain?*

Who's Who and What Do They Do?

To assure patient safety, every healthcare worker must, on a reoccurring basis, be credentialed by an array of professional, State and Federal agencies.

Expensive and time consuming: Credentialing costs the U.S. healthcare system billions of dollars per year and it is time consuming. Credentialing one physician takes, on average, 100 days; a time period where that physician cannot practice.

Thanks to encrypted digital ledgers, mobile technology, and cloud services, this seemingly intractable bureaucratic nightmare is being reimagined and rebuilt as a high-speed, on-demand service able to support existing regulatory and statutory obligations at scale – improving patient safety and increasing healthcare professional availability.

If this technology can be trusted to credential hundreds of thousands of mobile healthcare professionals – what would it take to credential and authenticate millions of continuously learning medical devices?

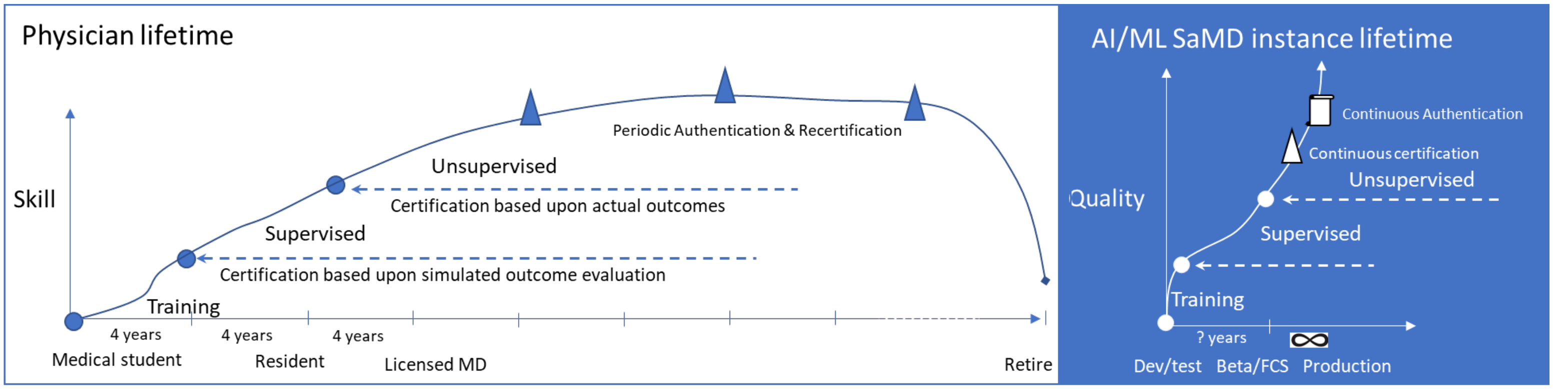


Figure C1: modeling an individual SaMD instance Quality as an independent healthcare worker's Skill.

The training, testing, and certification of a physician is not unlike the (ML)DLC or the FDA's GMLP for an AI/ML SaMD. The two only truly diverge after "certification." A physician is expected to continue to learn and improve – often in ways that are distinct from other physicians who were part of the same graduating class, a.k.a. the same release.

While there are governing bodies and controls in place to monitor the maturation of each individual physician – and to remove their privileges when needed – AI/ML SaMDs cannot be monitored individually today. As such, to assure patient safety, individual SaMD instance continued growth cannot be permitted.

Could a similar technology cocktail of encrypted digital ledgers (blockchain), mobile, and cloud technologies scale to reliably authenticate and then certify each individual medical device instance?

The first question that needs to be asked and answered is what innovation or benefits will be lost if continuous learning in production cannot be deployed. If there is no compelling use case, subsequent issues around monitoring and regulating their safety are moot.

What is evident is that, in order to remain relevant and support innovation, every interested party must remain open to reimagining the traditional roles and relationships between innovators, regulators, patients, service providers, et al alongside the coming waves of ML discoveries and breakthroughs.

ConnectedHealth

Advancing Transparency for Artificial Intelligence in the Healthcare Ecosystem

OCTOBER 2021

Connected Health is an initiative of ACT | The App Association

1401 K Street NW Suite 501, Washington, DC 20005
202.331.2130 | connectedhi.com

#connectedhealth

/ConnectedHealthInitiative

Executive Summary

Today, the most well-known FDA-approved applications of artificial intelligence and machine learning (AI/ML) technology in healthcare are diagnostic tools that help clinicians read and interpret images to predict, detect, and monitor a number of diseases, including diabetic retinopathy and lung cancer. In the future, the use of AI/ML technology in both operational and clinical settings promises to enable a more proactive approach to healthcare that promotes investments in preventative care that can result in fewer hospitalizations, fewer doctor visits, and fewer treatments. Across use cases, AI/ML technology is helping, and must increasingly help, the healthcare industry move away from a reactive disease treatment approach to a population health management approach that lowers costs and improves care.

The immense potential of AI/ML technology in healthcare may never be fully acheived, however, unless AI/ML technologies first earn the trust of healthcare professionals and patients. The cornerstone of building trust in AI/ML technologies is to enhance transparency – providing sufficient and appropriate information about the AI/ML, including its intended use, development, performance, and, when available, logic. The more understandable the decision-making process is for each individual technology, the more confidence there will be in AI/ML use in the healthcare system.

The recommendations in this Connected Health Initiative (CHI) AI Task Force report, informed by a public roundtable CHI held to address AI/ML transparency and extensive consultations with stakeholders from across the digital health ecosystem, represent a holistic approach to creating and maintaining the trust of both healthcare professionals and patients. The Task Force set out the foundational steps AI/ML tool developers must take to build transparency into their products, but it also outlines the important roles that clinicians, healthcare providers, regulators, academic medical institutions, and accrediting organizations must play.

The medical and technology communities have a shared responsibility to provide caregivers and patients (as well as other stakeholders) with an assurance of quality through truthful representations clearly indicating the AI/ML's intended uses and risks that would be reasonably understood by those intended and expected to use the AI/ML. Uptake will depend on the buy-in of clinicians who first develop trust in AI/ML software as a medical device (SaMD) through use and experience, establishing confidence as it is adopted into practice. Once adopted, clinicians can then work with their patients to explain their use of SaMD AI/ML and inspire the same trust and confidence from the patients in the output of the SaMD AI. Each step in this chain requires buy-in and support from policymakers (both within and outside of government).

The foundation of any successful use of AI/ML technologies in healthcare depends on the trust of healthcare professionals and patients, and we believe these recommendations present a clear path toward earning that trust.

ConnectedHealth

About the Connected Health Initiative

CHI is the leading multistakeholder policy and legal advocacy effort driven by a consensus of stakeholders from across the connected health ecosystem. We aim to realize an environment where Americans can improve their health through policies that allow for connected health technologies to enhance health outcomes and reduce costs. Having members who are developers and users of connected health technologies across a wide range of use cases, CHI serves as an active advocate before Congress, numerous U.S. federal agencies, and state legislatures and agencies. We seek to advance responsible pro-digital health policies and laws in areas including reimbursement and payment, privacy and security, effectiveness, and quality assurance, U.S. Food and Drug Administration (FDA) regulation of digital health, health data interoperability, and the rising role of artificial intelligence and machine learning (AI/ML) in care delivery.

In 2019, CHI formed a Task Force focused on policy challenges and opportunities related to the use of AI/ML in healthcare. CHI's AI/ML Task Force already developed a set of health AI/ML policy principles addressing how policy frameworks should adopt the role of AI/ML in healthcare.[1] A cornerstone of these principles is the idea of requiring those developing, offering, or testing healthcare AI/ML systems to provide truthful representations clearly indicating the intended use and risks that would be reasonably understood by those intended and expected to use the AI/ML solution. Such steps will provide much-needed quality assurances to caregivers and patients (as well as other stakeholders) and assist in resolving data issues that arise when an algorithm is fed bad data that can skew its learning and introduce bias. CHI's AI Task Force later developed detailed Good Machine Learning Practices for FDA-regulated AI,[2] which reflect and elaborate on this priority. The recommendations in this paper build on those deliverables.

Numerous CHI Steering Committee members and other key stakeholders from throughout the healthcare value chain participate in this Task Force and share a commitment to realizing the value of AI/ML in healthcare while protecting patient safety and advancing the quadruple aim. The recommendations in this paper find basis in an evaluation by the Task Force of the healthcare ecosystem's implementation of AI/ML to date, challenges and opportunities reflected by federal policymakers, and the existing and emerging issues created by AI's deployment. This report is also informed by a CHI public roundtable held in April 2021 on how to improve AI/ML transparency for caregivers and patients based on their needs and concerns, during which a wide range of stakeholders contributed to a discussion exploring novel approaches to transparency of AI/ML taken today.

For more information, please visit www.connectedhi.com.

[1] https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf.

[2] https://bit.ly/3B6nslm.

Artificial Intelligence's Role in a Successful Healthcare Ecosystem Requires Transparency

Responsible implementation of AI/ML in healthcare leads to improved medical outcomes and overall increased cost savings

Today, there are many important operational and clinical AI/ML solutions in use and many more in development.[3] Some of the most well-known applications of AI/ML in healthcare that have received market clearance from the FDA are diagnostic tools that help clinicians read and interpret images. For example, AI/ML image analysis software can assist clinicians in predicting, detecting, and monitoring a number of diseases, including diabetic retinopathy, lung cancer, prostate cancer, and skin cancer. Such AI/ML uses are generally intended to be used to assist human clinicians in providing more efficient and accurate results, rather than autonomously diagnosing disease.

Separately, research projects within and outside of clinical settings continue to further explore AI's potential to revolutionize healthcare. For example, an AI/ML system developed by researchers at Northwestern University's Feinberg School of Medicine correctly identifies small lung cancer tumors nearly 95 percent of the time, while radiologists undertaking the same task unassisted are correct only 65 percent of the time.[4] Researchers at Carnegie Mellon developed a miniature mobile robot called HeartLander that uses machine learning algorithms to make treating ventricular fibrillation (VF)—a deadly type of cardiac arrhythmia that requires cardioversion and then, if the patient survives, surgical removal of faulty heart tissue—far safer and less invasive.[5]

As a recent research paper discussing challenges related to deployment of AI/ML technologies into the clinical setting stated, "the success of a deep learning model does not rest solely on its accuracy." [6]The researchers noted that clinician "experiences with the system, and the socio-environmental factors that impacted system performance" must be evaulated and addressed for these systems to function in the clinical setting with the accuracy rates illustrated in the lab setting.[7] Clearly, if the challenges of integrating AI/ML tools into clinical workflow can be overcome, AI/ML can support clinicians in a wide range of other areas. Its potential to reshape the healthcare landscape is profound, especially in the improvements it can bring to any process within healthcare operation and delivery.

Medical devices and systems that use AI/ML also represent a real opportunity to drive down healthcare costs for consumers, practitioners, and healthcare businesses alike. It is estimated that AI/ML applications can cut annual U.S. healthcare costs by $150 billion by 2026.[8] Most of these cost reductions stem from changing the healthcare model from a reactive to a proactive approach, focusing on health management rather than disease treatment. This focus on using AI/ML as an investment in

[3] The FDA now publicly lists AI/ML medical devices cleared for marketing in United States, and includes their intended uses. *See* https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.

[4] https://www.nature.com/articles/d41586-020-03157-9

[5] https://onlinelibrary.wiley.com/doi/10.1002/rcs.2297

[6] Emma Beede et al, A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy, CHI Conference on Human Factors in Computing Systems (April 2020) available at https://dl.acm.org/doi/fullHtml/10.1145/3313831.3376718.

[7] *Id.*

[8] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7325854/.

preventative care can result in fewer hospitalizations, fewer doctor visits, fewer treatments, and thus fewer side effects. AI-based technology will have an important role in helping people stay healthy via remote monitoring technologies and coaching and will ensure earlier diagnosis, tailored treatments, and more efficient follow-ups.[9]

For example, AI/ML image analysis technologies can reduce medical expenses in several ways. For one, AI/ML systems can be very helpful in augmenting a clinician's analysis and treatment decisions more quickly. AI/ML technologies enable clinicians to provide the same, accurate service in a fraction of the time, increasing the volume of patients without increasing time spent treating them.[10] Second, a patient whose disease is diagnosed early will pay less to treat or cure the disease than one who catches it later. The longer a disease goes undiagnosed, the more damage it causes and more resources it takes to treat, assuming it remains treatable at all. Wearable technologies that use AI, such as remote monitoring technologies, increase access to healthcare and increase engagement in treatment plans by, for example, analyzing user health data in real time and notifying wearers or their healthcare providers (or both) of potential health issues.

By introducing new, accurate, and timely data streams for human clinicians' review, AI/ML medical tools and systems that use wearable technologies can enable practitioners to come up with care and treatment options without having to see a patient in person as much, reducing administrative and in-office visit resource expenditures, and, during outbreaks of communicable diseases, at lower risk of infection to both provider and patient. The use of such technologies will also enhance patient engagement in their own care plans. This same concept also applies to laboratory technologies that use AI/ML systems, where the work hours currently required for repetitive and routine tasks could see drastic reductions, significantly cutting labor costs.[11]

Increased efficiency, precision, and affordability are just some of the benefits that AI/ML can offer the healthcare community and those they serve, but realizing these benefits will depend on the buy-in of the provider and patient communities as well as support for responsible deployments from policymakers. CHI's AI/ML Task Force released detailed policy principles,[12] as well as proposed good machine learning practices for AI/ML meeting the definition of a medical device,[13] to address these challenges. Notably, CHI's AI/ML Task Force has acknowledged that without its processes being understandable by humans and transparency (providing sufficient and appropriate information about the AI/ML, including its intended use, development, performance, and, when available, logic), particularly for patients and caregivers, AI/ML cannot most effectively improve healthcare. Namely, those developing, offering, or testing healthcare AI/ML systems must provide truthful and understandable representations regarding intended use and risks that would be reasonably understood by those intended, as well as expected, to use the AI/ML software as a medical device (SaMD) solution.

[9] *Id.*

[10] *See* McPhail et al, Stage at diagnosis and early mortality from cancer in England (Br J Cancer 2015), doi: 10.1038/bjc.2015.49.

[11] Rong, et al, "Artificial Intelligence in Healthcare: Review and Prediction Case Studies," Engineering, doi: 10.1016/j.eng.2019.08.015 at Sec. 2.2.

[12] https://actonline.org/wp-content/uploads/Policy-Principles-for-AI.pdf.

[13] https://bit.ly/3B6nslm.

How Can Transparency into Healthcare AI/ML Solutions be Advanced?

While evidence of healthcare AI's potential for widespread benefit continues to build, that potential can never be realized without healthcare professionals and patients understanding and trusting AI/ML solutions. The more transparent the decision-making process is for each individual technology, the more confidence there will be in AI/ML use in the healthcare system.[14] Transparency for healthcare AI's intended uses must happen at several levels, disseminating tailored messaging to specific audiences that require insights into the AI/ML solution to make informed decisions. Building the trust that must be a foundation for the responsible deployment of AI/ML is a shared responsibility amongst developers, providers, and regulators.

Providing transparency into health AI/ML must start with the developers of the AI/ML tools. Then, uptake of AI/ML will need to be built on the buy-in of clinicians who first develop trust in AI/ML SaMD through use and experience, establishing confidence as it is adopted into practice. Once adopted, the provider can then work with his or her patients to explain their use of SaMD AI/ML and inspire the same trust and confidence by the patient in the output of the SaMD AI. Each step in this chain requires buy-in and support from policymakers (both within and outside of government).

The CHI AI/ML Task Force's recommendations for enhancing transparency for health AI/ML include:

Developers of AI/ML SaMD should**:**

- Prioritize making healthcare AI/ML solutions reasonably safe, efficacious, and equitable from the earliest stages of design, considering the perspectives of both patients and providers, leveraging and where necessary tweaking medical AI/ML guidelines on research and ethics,[15] leading standards,[16] and other resources as appropriate.
- Employ algorithms that produce repeatable results and, when feasible, are auditable, and make decisions that, when applied to medical care (such as screening, diagnosis, or treatment), are clinically validated and where possible understandable using rigorous procedures with documented methods and results, fostering efficacy through continuous monitoring.
- Rigorously identify, disclose, and mitigate biases in datasets used to train algorithms.
- Utilize risk-scaled privacy protection mechanisms for patients' data to account for the fact that the analysis by health AI/ML tools provides greater potential utility of those data items to other individuals, entities, and machines, providing many new uses for, and ways to analyze, the collected data, as well as correspondingly stronger incentives for malefactors to attempt to obtain access unlawfully. Specific uses of data that require additional safeguards (such as genomic

[14] https://www.bsigroup.com/globalassets/localfiles/en-gb/about-bsi/nsb/innovation/mhra-ai-paper-2019.pdf

[15] *E.g.*, World Health Organization, 'Ethics & Governance of Artificial Intelligence for Health' (2021), *available at* https://www.who.int/publications/i/item/9789240029200.

[16] *E.g.*, Consumer Technology Association, 'The Use of Artificial Intelligence in Health Care: Trustworthiness (ANSI/CTA-2090)' (2021), *available at* https://shop.cta.tech/collections/standards/products/the-use-of-artificial-intelligence-in-healthcare-trustworthiness-cta-2090.

information) may necessitate a tailored approach or enhanced protections from discrimination (e.g., pre-existing conditions or genomic information may be needed for patients).

- Comply with all applicable legal and regulatory requirements.
- Develop a tailored communications and engagement plan that gives patients and providers representative of the AI/ML tool's user group a reasonably justifiable level of confidence in healthcare AI's efficacy. Such communications should enable these patients and providers to visualize the AI, and to receive direct and clear information about how their health data are being collected and used (while also avoiding information overload) and how biases in data that exacerbate disparities in healthcare are being mitigated. Reflecting that the division of labor between the developers of AI-enabled tools and the clinician or patient is critical, clearly explain intended uses, including whether a tool might include the restriction that it is not for diagnostic use or for informational purposes only, as well as risks.

Providers should**:**

- Develop their own risk-based and tailored communications and engagement plan that enables them to explain to patients the development of the AI/ML application, its maintainnace, its performance, and how it aligns with the latest best practices and regulatory requirements to improve patient safety using easily understood and standardized formats. Providers should also acknowledge that "best practices" are dynamic and prone to obsolescence.
- Offer further detail for patients in additional resources that explain the clinical testing of AI/ML applications and the confirmation of the results by clinical experts.

The Food and Drug Administration (FDA) should**:**

- Leverage its successful approach to authorizing medical device AI[17] that has already safely brought health AI/ML innovations to patients and providers to develop a comprehensive regulatory approach to AI/ML that meets the definition of a medical device. The FDA can accomplish this by, for example, progressing its Software Precertification Pilot[18] to a full program available to all developers of SaMD AI, FDA can also update its rules and processes to realize its envisioned total product lifecycle (TPLC) regulatory approach, facilitating a potentially rapid cycle of product improvement and allowing these devices to continually improve while providing effective safeguards. This new approach should leverage CHI's Good Machine Learning Practices to address both locked and continuously learning AI.
- Evolve its requirements on reporting type and frequency so that such requirements can be adapted and scaled based on relevant factors such as risk, extent, and magnitude of

[17] Software as a Medical Device (SaMD): Clinical Evaluation: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm524904.pdf; Deciding When to Submit a 510(k) for a Software Change to an Existing Device: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm514737.pdf.

[18] Pre-Cert Program Version 1.0 Working Model: https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/UCM629276.pdf.

modifications, and the demonstrated reliability of the AI (e.g., quality control plans for updates).[19] Initially, the FDA should finalize guidance on SaMD pre-specifications and algorithm change protocol inputs that FDA should periodically receive.

- Develop methods to efficiently communicate when FDA has authorized a product developed with or that utilizes AI/ML, along with information on how it was developed, is maintained and performs, and aligns with the latest best practices and regulatory requirements that ensure patient safety using easily understood (e.g., infographics) and standardized formats. For example, where approval is required for the deployment of new solutions in the market, the FDA should provide information describing the datasets used to train the AI/ML software and what efforts are being taken to align with ethical standards and to mitigate data biases. This work should build on the recently released database of AI-enabled devices legally marketed in the United States from the FDA's Digital Health Center of Excellence.[20]
- Serve as a coordinator and convenor of other U.S. federal agencies to ensure a harmonized approach to health AI/ML transparency across government.
- Build on its leadership to date within the International Medical Device Regulatory Forum (IMDRF), promote its approach to SaMD AI/ML to improve approaches to transparency internationally.
- Host recurring public events, in partnership with health AI/ML developers, patients, and providers, that feature the FDA Digital Health Center of Excellence's latest approaches and thinking, as well as demonstrations of AI/ML in healthcare today.

The Centers for Medicare and Medicaid Services (CMS) should:

- Continue to develop its understanding of medical AI/ML definitions, present-day and future AI/ML solutions, how AI/ML is changing the practice of medicine, and the future of AI/ML medical coding.
- Develop Medicare support mechanisms for the use of AI/ML by providers based on clinical validation, alignment with clinical decision-making processes familiar to providers, and high-quality clinical evidence.
- Build on support provided in the Medicare system for the use of health AI,[21] develop easy to understand resources for Medicare beneficiaries that capture how AI/ML is being used in the Medicare system and what it means to patients. CMS should leverage its Advisory Panel on Outreach and Education[22] to develop this messaging.

[19] As the FDA has noted, new reporting mechanisms for a scalable AI/ML medical device reporting structure "may require additional statutory authority to implement fully". 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 (Apr. 10, 2021) at 15. *Available at* https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf.

[20] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices. This FDA list currently provides key information such as submission number, device and company name, and date of marketing authorization of the device (510(k) clearance, granting of De Novo, or PMA approval).

[21] For example, CMS already provides payment for CPT code 92229 (point-of-care diabetic retinopathy automated analysis and provides a diagnostic report using AI).

[22] https://www.cms.gov/Regulations-and-Guidance/Guidance/FACA/APOE.

The Federal Trade Commission (FTC) should:

- Support ways to mitigate biases or other unfair outcomes from healthcare AI,[23] and, where appropriate, enforce against violations of key laws such as Section 5 of the FTC Act, which prohibits unfair or deceptive practices, where appropriate.

Accrediting and Licensing Bodies, and Medical Specialty Societies and Boards should:

- Develop medical standard of care and ethical guidelines to address emerging issues with the use of SaMD AI/ML in healthcare needed to advance the quadruple aim.
- Develop and disseminate guidance and education on the responsible deployment of SaMD AI, both generally and for specialty-specific uses.

Academic and Medical Education Institutions should:

- Develop and include curriculum that will advance understanding of and ability to use healthcare AI/ML solutions, which should be assisted by inclusion of non-clinicians, such as data scientists and engineers, as instructors. Ongoing training and continuing education should also advance understanding of the safe and effective use of AI/ML in healthcare delivery, addressing both its capabilities and limitations.
- Develop curriculum to advance understanding of data science research to help inform ethical bodies such as Institutional Review Boards (IRBs) that are reviewing protocols of clinical trials of AI-enabled medical devices.

[23] https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai

Conclusion

CHI is pleased to present its recommendations on AI/ML transparency for the consideration of the healthcare ecosystem, policymakers, and others. We are committed to continued engagement with the digital health community writ large to realize the both the responsible deployment of AI/ML across healthcare and its immensely positive societal benefit.