May 7, 2024

The Honorable Ami Bera, M.D.
U.S. House of Representatives
172 Cannon House Office Building
Washington, DC 20515

Dear Representative Bera:

On behalf of the more than 90,000 members of the American College of Surgeons (ACS), I appreciate the opportunity to respond to your request for information on the state of Artificial Intelligence (AI) in health care. The ACS is dedicated to improving the care of the surgical patient and to safeguarding standards of care in an optimal and ethical practice environment. As such, we understand the critical role that technology plays in achieving this mission, as well as the need for thoughtful policymaking to ensure that tools such as AI are used with the utmost regard for patients' rights and safety. As we discuss below, physician and patient trust in AI technology is crucial. It is essential that AI tools are trained and maintained with high quality, diverse, valid, and representative data; are regularly assessed for continued accuracy and reliability; that regulators engage clinical experts in the assessment of AI health tools; and that physicians' clinical judgement remains paramount.

The ACS appreciates Congress' attention to this critical issue and welcomes the opportunity to share our response to a few of the questions posed on the state of AI in health care.

Implementation

What areas of health care are benefiting the most from AI integration, and what are the primary challenges hindering further adoption?

AI is already being implemented in health care settings across the country, and its use is only going to accelerate. Physicians and care teams must intake, process, utilize, and communicate more information than ever before. AI-based tools present a tremendous opportunity to support physicians in organizing and managing this knowledge to be applied for improved patient care and reductions in administrative burden and physicians' cognitive load. Some areas where it has been particularly beneficial to date include medical imaging, risk calculation, medication management, clinical decision support, and administrative tasks like billing and scheduling. In order to realize the full potential of AI in health care, policymakers must ensure both physician and patient trust in AI technology. This will require a

standardized regulatory framework developed in collaboration with stakeholders possessing clinical and technical expertise that guides the development and validation of algorithms.

Validation of digital health tools, including AI applications, is truly essential to physician trust, improving care delivery, and avoiding patient harm. There are many aspects to validation. Validation is necessary in terms of the technology/algorithm used, sources of data and the data used to train the algorithm, whether the outputs are accurate and unbiased, and whether the tool is appropriate for the specific setting in which it is used. It is also important that there is a framework to ensure validity of the tool after implementation. AI-based tools and their outputs must be monitored over time to ensure validity as the tool learns and iterates. Federal regulators should work in collaboration with an appropriate specialty society, clinical expert, or physician informaticist to reinforce physician trust in the tool. Applicability and validation of digital health tools are two of the most critical areas for physicians to successfully realize the potential of these technologies.

Ultimately, digital health tools should reduce, not add to, a physician's cognitive burden. AI technology can enhance a physician's ability to gather, process, and exchange knowledge and ultimately improve patient care when the tool is developed using semantic data exchange standards in alignment with validated clinical workflows. This enables these tools to provide the right information to the care team at the right time and seamless incorporation into the clinical workflow.

Efficacy, Accuracy, and Transparency

What best practices are recommended to ensure sufficient availability and use of health data for AI-driven health care solutions?

The ACS supports efforts to expand the use of real-world evidence (RWE) in the development and maintenance of medical technology. RWE is clinical evidence regarding the use and the potential benefits or risks of a medical product derived from analysis of real-world data (RWD), data related to a patient's health status or delivery of care that can be collected from a variety of sources such as mobile devices, wearables, and sensors; patient generated data used in home-use settings; product and disease registries; claims and billing activities; electronic health records, and more. Such data can complement data that are collected through traditional means and enhance clinical decision-making.

For the Food and Drug Administration (FDA) and other regulators, RWE is necessary for monitoring the safety of drugs, devices, and emerging technologies such as AI. As devices that use AI evolve, RWD will be reported back to the FDA regarding the product's safety, effectiveness, and potential risks. The true power of AI-based software lies in its ability to improve over time instead of remaining static.

But this is problematic for regulation because the device that was approved or cleared may no longer be operating in a similar fashion as it learns. RWD is necessary to show that the AI-based device still functions appropriately and in the way that it was intended. RWD is also important for accurately training AI algorithms. These data should be high quality, diverse, valid, and representative of the uses for which it will be applied.

What guardrails or accountability mechanisms could be set to ensure end-to-end transparency?

Transparency is essential to both physician and patient trust in AI technology. First, there must be transparent documentation of AI algorithms' development and ongoing validation processes to ensure that AI products are safe and effective as they iterate over time. This should include transparency in the source of the data, or knowledge base, that was used to develop, train, and test the algorithms. Knowledge bases vary in quality and trust. The internet would be an example of a low quality, low trust knowledge base, while a clinical data registry, such as the ACS National Surgical Quality Improvement Program database would be a high quality, high trust knowledge base. When an AI-based tool is trained on a knowledge base with high rigor, it can be assumed that there is lower risk of error because the data used in development is high quality and trustworthy.

Second, there must be transparency regarding when and how AI output is applied to patient care. This could come in the form of a "watermark" that confirms that an AI-based product or decision is in line with the highest clinical, quality, and regulatory standards. Groups like the ACS would be well-positioned to provide such validation for surgery. Finally, there must be transparency establishing clear lines of responsibility and liability for AI usage. These guardrails should be established through a standardized regulatory framework in collaboration with clinical stakeholders.

How can we ensure guardrails are put in place to mitigate risks such as disparate impact from racial, ethnic, and other biases?

It is critical to consider bias when designing, training, and using AI health tools. Various forms of bias based on race, ethnicity, gender, sexual orientation, socioeconomic status, and more can be perpetuated through the use of certain advanced digital health tools, especially those using AI. Bias can manifest in digital tools in various ways. For instance, if an AI algorithm is trained with data that fails to include all patient populations for which the tool is used, this would introduce inherent bias. Bias could also be unintentionally written into algorithms, leading to outputs that could have a biased impact on certain populations. The context in which the tool is used should also be considered when trying to avoid bias. If the tool were trained on a certain population for a specific purpose and is applied in a

different setting with a different patient population with varying risk factors, this could also result in bias.

Building a framework through collaboration with stakeholders possessing clinical and technical expertise that guides the development and validation of algorithms can assist in reducing bias if done with a high level of rigor. The framework could include a checklist with certain steps that developers would have to complete to ensure algorithms have gone through rigorous testing and validation. By following the processes and validation criteria set forth by the framework, developers can ensure that the algorithms are free of significant bias and will output accurate predictions. This type of framework coupled with external validation that utilizes data across various practice settings and demographics, can also be applied periodically following the implementation of the tool, to ensure that as the algorithms take in real-time data, they are still achieving a high-level of accuracy.

What are accountability mechanisms that can be put in place to ensure that there is an accurate spread of information?

As discussed above, the nature of generative AI is to draw conclusions and produce content based on the data it is trained on and the question it is being asked. Therefore, guardrails that guarantee the quality of the data input are critical to ensuring the accurate spread of information. These include clear standards for data quality and reliability, verifying the credibility and expertise of sources used in AI algorithms, providing transparent documentation of data sources and methodologies, enabling independent validation and peer review of AI algorithms, and fostering a culture of transparency, openness, and accountability among stakeholders involved in AI development and deployment.

Ethical and Regulatory Considerations

With the increasing reliance on AI in health care decision-making, what ethical and regulatory considerations need to be addressed to ensure patient safety, privacy, and equity?

As mentioned above, the use of RWE will be necessary for regulators to ensure that AI products are safe and effective as they iterate over time. Any regulatory framework should require that AI applications are assessed, maintained, and updated over their lifetime to ensure continued clinical safety and effectiveness, but also technological integrity. AI tools must be reviewed to make sure they are still valid, reliable, and accurate as they learn.

In addition, the ACS believes strongly that AI tools should never replace a physician's clinical judgment; rather, the goal of these and other digital health tools is to enhance physicians' knowledge and augment their cognitive efforts. Medical care relies not only on science, but on the capabilities of the care team, the local

resources, and the goals of the patient. Care is highly personalized and requires a physician-patient interface where the medical knowledge is contextualized and personalized in a trusted manner for each patient and physicians are empowered to make clinical decisions. As we assess AI applications, part of the assessment must evaluate the insertion of AI knowledge artifacts into a human workflow. It is the AI application's utility in the workflow that makes a difference in the informed nature of care, in the diagnosis, and in the treatment.

How can the use of AI in health care provide benefits while safeguarding patient privacy in clinical settings?

The FDA holds an important role in ensuring the safe and appropriate application of AI technology. Physicians can place greater trust in devices using digital technology if these devices have received FDA clearance or approval. FDA approval is also important for patient trust. Patients should know when they are receiving AI-informed care, that it comes from validated instruments, and that their privacy will be protected.

What regulations, policies, frameworks, and standards should entities utilizing AI adhere to, and what mechanisms are in place or should be in place to supervise and enforce them?

AI health tools must be both (1) clinically and (2) technologically sound. Validity, reliability, and accuracy are required on both levels. The ACS believes that clinical experts, such as physician informaticists, are best positioned to determine whether data used in AI applications are the best quality and the most appropriate from a clinical perspective, and to monitor the technology for clinical validity as it evolves over time. The FDA should engage advisory groups for clinical and technical excellence that are condition or programmatically defined with cross specialty expertise, in order to ensure an AI tool is reliable and valid on multiple levels.

In addition, physicians and specialty societies are well-equipped to assist the FDA as they consider what tools and/or information would be most useful in driving improvements and advancements in clinical care and the format in which the information should be expressed. Understanding where physicians see the benefits of AI in their practices is crucial to help build trust in the capabilities of the technology, leading to broader utilization. Likewise, understanding why physicians decide not to use or do not trust certain health technologies in their clinical practices would also be useful as regulators certify products for real-time use.

As discussed above, it is especially important to emphasize that the data used to train algorithms is critical to their validity and reliability. The data should be high quality, diverse, valid, and representative of the uses for which it will be applied. While the data used to train the AI-based tool is important, it is equally important

that up-to-date data are used to retrain such tools so that the algorithms themselves remain current, reliable, and valid.

At the facility level, institutions should have their own governance and structure for AI-based tools, including pathways for user education, feedback and timely responses to feedback as physicians have concerns or encounter issues. Liability risks and uncertainty about who is responsible for issues with certain algorithms, outputs, or user errors can hinder implementation of these tools. Before leveraging AI technology, institutions should be confident in the quality of the tool and its capabilities.

Other Considerations

Are there legislative measures that Congress can take to ensure access to safe, reliable AI healthcare services?

Congress has an important role in establishing the regulatory framework that will govern the use of AI in health care. As discussed above, this should be a collaborative effort between federal agencies, such as FDA, and outside stakeholders with clinical and technical expertise. Congress should consider providing the resources and infrastructure necessary to support this public private partnership, such as grants to entities wishing to engage in AI development and/or validation efforts and opportunities to convene stakeholders. Additionally, Congress could take steps to establish an ongoing knowledge repository that could house both standards and validation information as well as a free, open-source synthetic patient environment that could be used to test the clinical and technical aspects of any AI application.

Concluding Remarks

The ACS thanks Congress for its thoughtful attention to the use of AI technology in health care and looks forward to continuing to work with lawmakers on these important issues. For questions or additional information, please contact Emma Zimmerman with the ACS Division of Advocacy and Health Policy at ezimmerman@facs.org.

Sincerely,

Patricia L. Turner, MD, MBA, FACS
Executive Director & CEO