December 22, 2020

Dr. Leith States
Chief Medical Officer
Office of the Assistant Secretary for Health
Office of the Secretary
U.S. Department of Health and Human Services
200 Independence Avenue, SW, Room 716G
Washington, DC 20201

Subject: (85 FR 73280; 2020-25328) Request for Information-Landscape Analysis To Leverage Novel Technologies for Chronic Disease Management for Aging Underserved Populations; Comments of the American College of Radiology

The American College of Radiology (ACR)—a professional association representing nearly 40,000 diagnostic radiologists, interventional radiologists, nuclear medicine physicians, radiation oncologists, and medical physicists—appreciates the opportunity to provide comments on the request for information (RFI) from the U.S. Department of Health and Human Services (HHS) Office of the Assistant Secretary for Health titled, "Landscape Analysis To Leverage Novel Technologies for Chronic Disease Management for Aging Underserved Populations," published in the Nov. 17, 2020 *Federal Register*.

ACR's Data Science Institute (DSI) empowers the advancement, validation, and implementation of artificial intelligence (AI) in medical imaging for the benefit of patients, society, and the radiological professions. To that end, the ACR DSI leverages the value of radiology professionals and radiology informatics experts through the development and maintenance of an open library of clinically relevant AI use cases and workflow integration resources.[1] It works to protect patient safety through algorithm validation[2] and longitudinal performance monitoring registries,[3] as well as by informing the regulatory policies of federal agencies to advance innovation while ensuring algorithm performance. Finally, the ACR DSI educates radiology professionals to demystify imaging AI, and works to make these tools accessible to all radiologists regardless of practice size, geographical location, or the patient populations they serve.

A. Barriers and Opportunities for Technology-Driven Solutions

5. How will training data sets be established and implemented to drive effective technology solutions that improve chronic disease outcomes for aging populations in rural areas?

The ACR DSI initiatives include programs such AI-LAB[4] without charge to facilities regardless of size or location. The AI-LAB platform allows facilities to build and/or evaluate AI models using their own patient population data at their own sites. Since AI algorithms can be very brittle if used on populations that differ from the data on which they were trained, it is important that AI models are trained using data from multiple institutions with varying patient demographics and imaging equipment. These algorithms should then be tested against an equally diverse multicenter data set, and finally, facilities that have not been part of the training process should evaluate AI models prior to deployment. The ACR AI-LAB

[1] ACR DSI. Define-AI. https://www.acrdsi.org/DSI-Services/Define-AI
[2] ACR DSI. Certify-AI. https://www.acrdsi.org/DSI-Services/Certify-AI
[3] ACR DSI. Assess-AI. https://www.acrdsi.org/DSI-Services/Assess-AI
[4] ACR DSI. ACR AI-LAB. https://ailab.acr.org

enables federated learning where an algorithm initially developed at one institution can be passed from institution to institution to train the algorithm with a broad array of patient data, and since any practice can use AI-LAB, developers can include data from the aging rural populations seen in the small and rural practices where the models will be used. The combination of federated learning and distributed testing and evaluation should enhance the development of more robust and accurate algorithms that are generalizable to at-risk populations.

As AI becomes more commonly deployed in clinical practice, physicians will also need educational resources about AI. The ACR AI AI-LAB also provides a path for hands-on learning and simulation to educate radiologists previously unfamiliar with AI and data science so they can be ready to use AI to improve patient care.

6. How will AI solutions be validated? What metrics will be used to evaluate the effectiveness of AI/machine learning algorithms?

To ensure validity across diverse populations, datasets used for validation of AI algorithms should be from multiple institutions with diverse patient demographics and disparate input devices (e.g., imaging devices and models) and should include patient data from the target populations. To ensure at-risk aging rural populations are well represented, the validation datasets could be developed in the same manner as federated training datasets whereby facilities that include the target populations can participate in creating the validation dataset and participate in the validation process without having to move sensitive patient information off site. The ACR AI-LAB platform provides the ability to locally annotate data with ground truth, such as expert evaluation of pathology concordance. The ACR DSI expert panel-defined AI use cases identify common data elements used in the criteria of AI performance evaluations (e.g., immediately comparing AI outputs to expert panel-defined ground truth) as well as metrics for model analysis. The evaluation metrics used to evaluate an AI algorithm depend on the type of AI use case. Specifically, for use cases with classification outputs, the algorithms will be evaluated using the confusion matrix comparing the algorithm's predictions and the ground truths, the accuracy, and the linear and quadratic Kappa score. Additionally, for binary classification use cases (for example, pneumonia absent or pneumonia present), the AUC (Area Under The Curve) ROC (Receiver Operating Characteristics) curve, sensitivity, and specificity will also be calculated for evaluation. For use cases with continuous outputs, the mean squared error, the mean absolute error, and the Pearson correlation coefficient will be used for evaluation. Finally, for use cases with volumetric outputs, like bounding boxes or regions of interest, the algorithms will be evaluated using the mean Dice coefficient, the mean intersection over union, and the mean average precision.

Additionally, AI models may degrade over time due to changes in imaging equipment or changes in the patient population; therefore, there needs to be a mechanism for real-world longitudinal performance monitoring of AI models once they are deployed in clinical practice. The ACR Assess-AI registry captures data to measure algorithm performance based on radiologists' interpretations as well as metadata about the examination so that if an algorithm breaks, the potential causes can be analyzed. Imaging AI effectiveness can also be evaluated over extended periods through ongoing monitoring of patient outcomes, particularly where algorithm findings are ancillary to reasons for ordered radiology studies.

7. How will healthcare team and patient trust in technology solutions be addressed? How will legal and ethical issues be addressed for technology solutions designed for improving chronic disease outcomes?

Trustworthiness of AI can only be achieved if it is built using structure use cases with specific data elements that define good machine learning practices. AI must integrate seamlessly into the clinical workflow. End users should expect and be able to verify the algorithm was validated using diverse data and then be able to evaluate the model in their own practice prior to use. It is important for AI-enabled solutions to be sufficiently generalizable for all situations in which these tools may be used, including with diverse input devices (e.g., different medical imaging devices and models) and different patient populations, with clinical validation and ongoing real-world performance monitoring conducted by

trusted entities. To that end, the ACR AI-LAB provides end users a platform to independently evaluate imaging AI algorithms prior to purchase, and the ACR Assess-AI registry provides information to end users about the performance of the model in their own practice over time. Provider- and patient-level education and demystification of AI is important for establishing trust in AI-enabled tools in general.[5]

8. How will bias and variance be addressed in machine learning algorithms for this application? How will supervised versus unsupervised learning be used to develop inferences and patterns from data sources? What will be the challenges and proposed solutions for data cleansing and transformation?

While some may believe that machines are unbiased, AI models are developed using parameters specified by humans and as such are potentially subject to unintended bias. We believe unintended bias can be mitigated through a combination of federated training and validation using data from a variety of institutions and then real-world performance monitoring that captures patient demographic data.[6] A federated approach can be used to ensure data from underserved and minority population groups are included in algorithm training. The AI-LAB platform allows for this federated approach, ensuring a sufficient diversity of data are included during training as well as validation.[7]

Vocabulary differences and unstructured narrative data are major challenges to the cleanliness and usability of data for AI/ML. The ACR and others have made significant progress on identifying common data elements and promoting structured reporting and imaging data interoperability.[8] Future AI innovations that use natural language processing to map human language to these common data elements from radiological reports may be quite beneficial.

B. Key Indicators & Data Sources of Technology-Driven Chronic Disease Management

3. What selected health conditions should be addressed as priority conditions to assess technology-driven capacity to influence access, timeliness, and quality of healthcare treatment and preventive services to aging populations living in rural areas?

Prioritization of technology-enabled chronic disease management should be informed primarily by the clinical needs of health care providers (i.e., usefulness of the technological solution in the real-world). The ACR observed early in the AI era that many novel imaging solutions focused on by researchers and the media were not clinically useful nor readily integrated into radiology workflows. To address this, the ACR DSI worked to define use cases for imaging AI and to make these freely available in a public directory to help drive innovators toward developing solutions of the most clinical need and utility.

D. Public-Private Partnerships

1. Provide ideas of the form and function of a public-private partnership model to leverage the adoption of technology-driven solutions to improve outcomes for at-risk populations such as aging Americans living in rural areas.

Most importantly, HHS should collaborate with national health care professional associations, such as the ACR, to ensure any identified technology-driven solutions would be implementable and useful in real-world practice before expending limited resources. Practice size is not a barrier to collaborations with radiologists serving unique at-risk populations such as aging Americans living in rural areas, and the ACR's Data Science Institute's toolkit allows these radiologists be part of the development, evaluation, and implementation of AI-enabled solutions. Physicians and other health care providers generally have good awareness of the needs and opportunities for technological augmentation of care. The ACR DSI

---

[5] European and North American Multisociety Statement, 2019. Ethics of AI in Radiology. https://www.acr.org/-/media/ACR/Files/Informatics/Ethics-of-AI-in-Radiology-European-and-North-American-Multisociety-Statement--6-13-2019.pdf

[6] Wood, M.J., 2020. American College of Radiology | Addressing Bias in Radiology. https://www.acr.org/Practice-Management-Quality-Informatics/ACR-Bulletin/Articles/September-2020/Addressing-Bias-in-AI

[7] Cohen-Addad, D, 2020. American College of Radiology | Democratizing AI. https://www.acr.org/Practice-Management-Quality-Informatics/ACR-Bulletin/Articles/December-2020/Democratizing-AI

[8] ACR DSI/ACR-Assist. Computer-Assisted Reporting/Decision Support Assistant for Radiologists. https://assist.acr.org/assistweb/overview

leverages expert panels of radiologists, informaticians, and patients to establish a public directory of clinically valuable imaging AI use cases, and this directory could be used by HHS agencies to guide prioritization of regulatory or research funding efforts related to radiology AI. The ACR DSI use case development process transforms human language to machine language and as such, if there are population specific ideas for where AI could play a role in improving care, the ACR is happy to discuss the opportunity to develop these ideas into AI use cases for model development.

2. What organizations, groups, and/or, associations should HHS engage as part of such a collaborative effort?
We encourage HHS to collaborate with the ACR DSI and leverage DSI's assorted programs and services. Likewise, other national physician associations inspired by the ACR DSI model have engaged in AI use case development and/or validation initiatives.

As always, the ACR welcomes the opportunity for further dialog and collaboration with the Office of the Assistant Secretary for Health and other HHS agencies. Please contact Gloria Romanelli, JD, ACR Senior Director of Legislative and Regulatory Relations, at gromanelli@acr.org, or Michael Peters, ACR Director of Legislative and Regulatory Affairs, at mpeters@acr.org or (202) 223-1670.

Sincerely,

Howard B. Fleishon, MD, MMM, FACR
Chair, Board of Chancellors
American College of Radiology