RADIOLOGICAL SOCIETY OF NORTH AMERICA
820 JORIE BLVD., SUITE 200, OAK BROOK, IL 60523-2251
TEL 1-630-571-2670 FAX 1-630-571-7837
RSNA.ORG

RSNA® Radiological Society of North America

March 14, 2024

Susan Gregurick, PhD
Associate Director for Data Science & Director of the Office of Data Science Strategy (ODSS)
National Institutes of Health (NIH)

Dear Dr. Gregurick,

The Radiological Society of North America (RSNA) is a non-profit organization representing over 48,000 medical imaging professionals in 31 radiology subspecialties from 160 countries around the world. Our mission is to promote excellence in patient care and health care delivery through education, research, and technological innovation.

Radiology and medical imaging are data intensive specialties at the forefront of grappling with the numerous ways that big data, artificial intelligence, and machine learning are transforming the practice of medicine and the delivery of healthcare. As a leading medical society with expertise in health informatics, data science, and the use of artificial intelligence (AI) in imaging and medical practice, RSNA appreciates the opportunity to provide comments in response to NIH's draft *Strategic Plan for Data Science 2023-2028*. Our comments below focus on Goals 1-3 of the draft Strategic Plan.

Goal 1: Improve Capabilities to Sustain the NIH Policy for Data Management and Sharing

RSNA supports the NIH's goals of developing sustainable data resources to support open research. Adherence to emerging data stewardship and sharing principles, including FAIR, is critical and RSNA applauds NIH's commitment to these principles. As the NIH seeks to implement Goal 1 of the draft Strategic Plan, RSNA stands ready to partner with the agency to develop platforms and training resources that bolster the ability of the biomedical research community to manage, share, and sustain FAIR data and strengthen the broader data repository and knowledgebase ecosystem.

RSNA, along with our partners at the University of Chicago, the American College of Radiology (ACR), and the American Association of Physicists in Medicine (AAPM) supports the NIBIB-funded Medical Imaging and Data Resource Center (MIDRC) project. Through its own data-gathering activities and its engagement in the MIDRC consortium, RSNA has continued to expand its capabilities in research data collection, annotation, curation, and publication. Since August 2020, RSNA has collected COVID-related imaging and clinical data for use in the imaging community's response to the pandemic. We have established a consortium of healthcare and research organizations able to contribute de-identified imaging studies and associated clinical information to this public data resource for AI research. To date, we have collected over 309,000 imaging studies and curated more than 130,000 of these studies for publication through the consortium. As the draft Strategic Plan notes, MIDRC is a prime example of NIH-supported biomedical data infrastructure that must be integrated into a federated data infrastructure, including the use of privacy-protecting identifiers to link subjects across multi-modal data repositories.

As part of its role in MIDRC, RSNA has assembled a team of dedicated imaging data curators and implemented a data ingestion, processing, and curation pipeline using a suite of tools for imaging data review and management. Since the early 2000s, RSNA has developed and maintained the Clinical Trial Processor (CTP) tool and licensed it freely to the research community. CTP has become the industry standard in medical imaging for data de-identification, processing, and transmission. RSNA is currently developing a new version of this suite of tools called 'RSNA Anonymizer' using contemporary programming methods (Python, continuous integration) designed to simplify adoption by a variety of radiology sites and to work more seamlessly with cloud-based data environments. RSNA Anonymizer will be released under an open-source license to enable community adoption and reuse. RSNA plans to expand the capabilities of its data deidentification tools in pixel masking (both for alphanumeric protected health information and for facial contours) and concept extraction from free-text radiology reports. RSNA has worked with AWS Professional Services to implement an access-controlled web front-end for RSNA's S3 data storage that enables secure upload for data contributing sites and access to submitted data for RSNA's current data processing tools. And, working with contracted partner MD.ai, RSNA has built a growing array of curation and quality-checking tools for imaging data curation, management, and annotation. These innovations have enabled RSNA to significantly streamline data submission and increase data throughput capacity.

Goal 2: Develop Programs to Enhance Human Derived Data for Research

RSNA supports NIH's goal of expanding the use and utility of real-world healthcare-derived data for research. As the draft strategic plan notes, bridging the gap between health care and research settings requires a shared set of standards, practices and tools for privacy protected data collection and exchange. RSNA's experience both in collecting and curating data for research and in developing standards to enhance data quality and accessibility positions us well to collaborate on initiatives that address this goal.

RSNA has a long history of developing foundational standards for data sharing in medical imaging. Since the initial demonstrations of the DICOM standards at the RSNA's Annual Meeting beginning in 1993, RSNA has advocated for the implementation of data standards to enable quality and consistency and the aggregation of data to support research and technological advancement in radiology and medical imaging across specialties (e.g. pathology; ophthalmology; etc.). RSNA oversaw the work of developing foundational terminology resources in medical imaging through the development of the RadLex ontology and the LOINC/RSNA Radiology Playbook or radiologic procedure names and codes. The World Health Organization (WHO) has recently initiated discussions with RSNA toward integrating RadLex with the WHO Family of International Classifications. Currently, we are working jointly with the American College of Radiology to expand a core set of common data elements (CDEs) for radiology in the RadElement.org CDE library, designed to provide improved uniformity in recording radiologists' diagnostic observations and making consistent data available for clinical use, outcomes analysis, and research.

As a founder of Integrating the Healthcare Enterprise (IHE), RSNA has fostered the creation and adoption of interoperability standards for workflow improvements and data sharing in radiology, as well as cardiology, dentistry, eye care, pathology and laboratory medicine, and radiation oncology. IHE's technical specifications include those underpinning the nationwide health information network infrastructure implemented under the Trusted Exchange Framework and Common Agreement (TEFCA) administered by the Office of the National Coordinator for Health IT (ONC). Specifications published by IHE in radiology include implementation guides for using standards including DICOM and HL7 FHIR to integrate AI-generated findings as discrete structured data elements in radiology reports.

With funding from NIBIB, RSNA implemented the pilot Image Share Network, enabling patient access to imaging records at 12 sites across the US. This included exploratory work with the then nascent *All of Us* Research Program to enable access of medical imaging through EHR data portals and a common process for patient consent to protected use of data in medical research. RSNA would be eager to partner with NIH to expand efforts to include medical imaging in the spectrum of clinical data made available for research through the *All of Us* program.

RSNA's work on the Image Share project included collaborating with Sequoia Project (currently the Recognized Coordinating Entity for ONC's implementation of TEFCA) to develop the Image Share Validation Testing Program to assess the standards conformity and standards-based image sharing capabilities of radiology systems. In 2019 RSNA also collaborated with Carequality (Sequoia Project's sister organization, which provides the interoperability framework linking more than 4,200 hospitals and 50,000 clinics nationwide) to publish the Imaging Data Exchange Implementation Guide, which extends current data sharing infrastructure to exchange more than 400,000,000 medical documents monthly to incorporate medical imaging data. RSNA would look forward to collaborating with NIH in enabling research data collection using the emerging US health information infrastructure and expanding the adoption and use of interoperability standards and standardized CDEs to support research.

RSNA appreciates the emphasis on cross-disciplinary training highlighted in the draft plan as a key strategy for achieving expanded use of clinical data in research. RSNA has worked to make informatics and data science integral components of training for radiologists and associated professionals. RSNA has partnered with the Society for Imaging Informatics in Medicine (SIIM) to present the National Imaging Informatics Course, a weeklong series of educational sessions held twice annually since 2016 designed to provide radiology trainees with essential training in informatics and data science. In 2022, RSNA launched its Imaging AI Certificate Program, a curriculum focused on imaging AI essentials for radiologists, including practical training on the implementation, maintenance and use of AI, data science and informatics, analytics, and ethics. RSNA stands ready to partner with the NIH on new approaches to cross-disciplinary training in imaging informatics and the use of AI in medical imaging.

Goal 3: Provide New Opportunities in Software, Computational Methods, and Artificial Intelligence

AI is redefining the practice of medicine, and perhaps no specialty has been more impacted by AI than radiology, which has seen greater development and application of AI algorithms than any other medical field. Of the almost 700 algorithms currently cleared for use by the FDA, more than 76% are for use in radiology. Hundreds of imaging AI algorithms are currently undergoing consideration for regulatory approval, entering the marketplace, and becoming part of clinical care, yet development and implementation of effective imaging AI tools is limited by the availability of quality data for the training and validation of AI algorithms, as well as by the radiologist's ability to annotate these training data, a time intensive process made more difficult by the lack of standardized tools.

RSNA leverages its leadership in medical imaging to coordinate and accelerate the implementation of AI in radiology and is viewed by the broader medical imaging community as a trusted source for data collection and annotation. RSNA's approach is informed by and aligned with the Food and Drug Administration's (FDA) Good Machine Learning Practice for Medical Device Development principles. We have responded to the current wave of innovations in AI for medical imaging with an extensive set of programs covering AI education, ethics and standards, communications and research. RSNA's efforts here are consistent with President Biden's Executive

Order on the *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, which calls for the development of "standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy." As the NIH seeks to implement Goal 3 of the draft Strategic Plan, RSNA stands ready to support the agency in the development of tools and training opportunities that help researchers create and prepare data that are FAIR and AI-ready; and in operationalizing processes and programs that leverage community engagement to generate and validate diverse, equitable, and inclusive data and methods for AI in medical imaging.

As the draft Strategic Plan recognizes, for AI advances to result in meaningful improvement in patient care, it is imperative that the community have access to large, diverse, high-quality datasets and reliable methods for model training and validation. RSNA has a longstanding role in collecting and publishing data to support AI research.

Since 2017, RSNA has conducted a continuing series of AI Challenges, collecting and annotating ground truth imaging datasets to address high-value clinical applications, foster research, technological innovation, and focus the AI research community on vital use cases. RSNA has conducted seven medical imaging AI Challenges to date, with our eighth challenge currently in progress. Anywhere from several hundred to well over a thousand teams of researchers from around the world have participated in each of these challenges. For many of these researchers, these competitions were their first opportunity to work in medical imaging AI and a number of the winning teams featured cross-disciplinary combinations of researchers, data scientists, and physicians. We have assembled datasets in multiple imaging modalities and addressing research areas such as determination of bone age, detection of cervical spine fractures, pneumonia, localization of intracranial hemorrhage, and segmentation of brain tumors. These datasets include tens of thousands of imaging studies (over 2.5 terabytes of data) with annotations provided by subspecialty expert radiologists and supporting clinical data. They are organized specifically to support AI research. RSNA has access to subject matter experts and the capabilities to organize and accurately annotate (including both segmentations and textual diagnostic labels) contributed data to create effective AI training and validation datasets. Importantly, these competitions are also valuable for AI tool developers seeking to ensure their products are trained on FAIR data.

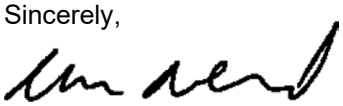
Consistent with RSNA's educational mission, the Society has facilitated the engagement of the broader AI research community by convening vendors and users and providing training materials and tutorials. The winning models developed under our AI Challenges have been released under open licenses, along with explanatory documentation. The datasets developed for the competitions have also been made persistently available to the research community under open licenses (with representative data samples reserved for model testing and validation). RSNA's Radiology AI Data Standards Committee has developed a standard template to describe the source attributes of imaging datasets and published this information along with the datasets. Finally, the task forces overseeing each of these challenges have published research articles and presentations at medical meetings describing the datasets and methodologies for collection, curation, and annotation, as well as the organization and outcomes of the competitions. More than as an education tool, this practice champions safe and transparent AI ensuring that downstream users of these models, both caregivers and patients, have the best possible chance of understanding how these models make inferences.

In collecting and curating data to support AI research, RSNA has actively sought to foster the development of ethical AI by ensuring diversity of health data used to train and validate algorithms. RSNA has leveraged its unique international reach within medical imaging to build a progressively broadening consortium of data-contributing research sites. For the Abdominal Trauma AI Challenge conducted in 2023, data were contributed by 23 different institutions in 14 countries across six continents. However, it is important to note that despite the

reach of RSNA's AI Challenges, important populations continue to be underrepresented in AI tool development, including pediatric populations and other groups historically underrepresented in biomedical research. RSNA has sought to make assessment of AI bias and fairness an integral and growing part of the AI research it supports, collaborating with researchers focusing on AI fairness in the laboratory of Marzyeh Ghaseemi at the Massachusetts Institute of Technology to analyze the performance of algorithms developed in the AI challenges it sponsors for fairness and detectable bias across different patient populations.

In conclusion, RSNA stands poised to support and partner with the NIH in its efforts to support data driven discovery through the generation, management, and dissemination of FAIR data. We believe our expertise in medical imaging informatics and the use of AI in imaging can assist the NIH in meeting many of the objectives included in the draft Strategic Plan. RSNA would welcome the opportunity to engage with the NIH Office of Data Science and Strategy for further follow-up discussions on any of the points raised above. For additional information or questions please contact RSNA's director of government relations, Libby O'Hare (eohare@rsna.org).

Sincerely,



Umar Mahmood, MD, PhD
Chair of the Board
Radiological Society of North America

cc: Pari V. Pandharipande, MD, MPH, RSNA Board Liaison for Government Relations
Pedram Heidari, MD, Chair, RSNA Government Relations Committee