Radiology: Artificial Intelligence

COMMENTARY

Artificial Intelligence in Radiology: The Computer's Helping Hand Needs Guidance

Evis Sala, MD, PhD • *Stephan Ursprung, MD*

Evis Sala, MD, PhD, is the professor of oncological imaging at the University of Cambridge, UK and co-leads the Advanced Cancer Imaging Programme and the Integrative Cancer Medicine Programme for the Cancer Research UK Cambridge Centre. Her current research focuses on radiogenomics through multiomics data integration for evaluation of spatial and temporal tumor heterogeneity and on the applications of AI methods for image reconstruction, segmentation, and data integration.

Stephan Ursprung, MD, is a 3rd-year PhD student in the department of radiology at the University of Cambridge, UK. His research focuses on the development of AI models for automated segmentation, lesion classification, and treatment response prediction in renal cancer. Dr Ursprung's interests include health information technology, molecular and physiologic imaging, as well as multiomics data integration.

Artificial intelligence (AI) is not entirely new to the medical field. We are accustomed to applying tools that have been developed with a varying degree of human and computer input in our clinical practice. Representative examples include handcrafted diagnostic algorithms to triage patients presenting with acute illness (1), statistically derived scores for osteoporotic fracture risk assessment (2), and decision trees for the differentiation of benign and malignant ovarian masses (3). Common to all these tools is that changes to input parameters lead to predictable changes in model output, making them easy to interrogate and understand. More sophisticated, deep learning (DL)–based models employed in decision support systems have been implemented in clinical practice for the automated interpretation of electrocardiograms (4) and detection and classification of lesions on mammography (5). Although DL-based models have exciting potential to solve complex problems, their black box approach faces skepticism despite advances in interpretability and explainability (6).

The recent, widespread availability of hardware and software for the development of AI solutions for medicine has inspired an exponential increase in publications. Data-rich medical specialties such as radiology have become a particular focus of rapid development. However, there are ample scope and encouraging initiatives for AI to support the delivery of care from general practice, primary care, the emergency department, and specialist diagnostics to patient self-care (7).

This study by Tadavarthi and colleagues (8) has examined the market of AI-enabled image analysis solutions for radiology and provides recommendations for the evaluation of AI tools before purchase. In their market study, the authors illustrate how most solutions are focused on high-volume conditions. Unsurprisingly, many solutions focus on support for lesion detection and quantification rather than decision support for diagnosis and recommendations for management where regulatory stakes and hurdles are higher. Yet only a minority of solutions advertised at the Radiological Society of North America and Society of Imaging Informatics in Medicine annual meetings between November 2016 and June 2019 have received approval for the American or European market. This finding is indicative of a rapidly developing field where, after years of purely scientific development, the first tools start undergoing consolidation, approval, and marketing. The sole focus on solutions advertised at North American conferences risks missing tools by smaller companies with lower marketing budgets and introducing a geographic bias. Indeed, several other solutions have achieved Conformité Européenne marking or U.S. Food and Drug Administration (FDA) approval. Such approval or their equivalent in other territories is a precondition for the implementation of products, but it is by no means sufficient to identify clinically beneficial and financially viable tools. Overall, the adoption of these algorithms into clinical practice is emerging, and further work is needed to transform the scientific enthusiasm for developing advanced AI tools with a broader scope into clinically workable solutions. For tracking the ongoing market, surveys like this study (dating from November 2019) date quickly, leaving a gap for a living review and other market watchers.

Tadavarthi et al contribute to a growing number of recommendations for the acquisition and adoption of AI solutions in medicine with a particular focus on radiology (9). They raise important considerations to determine whether an AI tool is a viable solution for an individual service. In addition, one might want to consider the following criteria:

From the Department of Radiology and Cancer Research UK Cambridge Centre, University of Cambridge, Box 218, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0QQ, England. Received August 31, 2020; revision requested September 9; revision received September 15; accepted September 28. **Address correspondence to** E.S. (e-mail: *es220@medschl.cam.ac.uk*).

Conflicts of interest are listed at the end of this article in this issue.

See also the article by Tadavarthi et al in this issue.

1. How will the solution integrate into the current clinical pathway?
    a) Rule-in/rule-out/triage of patients
    b) First reader/second reader
    c) Equality of use case and approved use
2. How well will the algorithm generalize?
    a) Comparability of patient demographics (age, sex, ethnicity) between testing dataset and intended patient cohort
    b) Comparability of the clinical setting (screening, diagnostic, investigative, or therapeutic setting) between testing dataset and intended use case
    c) Evidence from monitoring programs and ongoing trials
3. Does the solution have a clinical and/or cost-benefit in real use?
    a) Equivalent or improved diagnostic accuracy: sensitivity, specificity…
    b) Overall cost-benefit: clinician time, reducing or increasing additional tests, overdiagnosis and overtreatment of indolent findings

Key to the adoption of AI algorithms discussed in the article is the ongoing review of a model's performance before implementation and during use. Changing equipment, improving acquisition and reconstruction methods, and evolving patient demographics may change the diagnostic accuracy. The requirement for an ongoing performance review complicates the adoption of AI-enabled tools for smaller, nonacademic providers and may require collaboration with the manufacturer or other providers. Continuous learning is not currently permissible under FDA regulations, and drifts in input data may lead to an increasing bias in prediction over time, which can only be corrected once a new version gains approval. However, continuous learning risks leading to bias and poor performance if an existing algorithm is trained on new, low-quality, and poorly labeled data. Most worrying, the institution itself will not notice a decline in performance due to poor ground truth labels (10). Ultimately, the postmarket performance in the field will determine the clinical value of AI solutions. With the introduction in Europe of the Medical Devices Regulation, beginning in 2021, manufacturers must track postmarket performance, off-label use, and adverse events systematically.

Tadavarthi et al suggest acquiring licenses that would allow expanding the application of models beyond their intended use. This advice is potentially unsafe outside an appropriate research setting. Researchers and clinicians should seek advice on the legal aspects of off-label use and investigational use permissions from regulatory authorities and ethics commissions where appropriate. Off-label use may result in unpredictable outputs and, as with algorithm validation and performance monitoring, validation for another indication requires a large, sufficiently diverse dataset to start with.

The authors recognize that the adoption of AI solutions in radiology is determined largely by the economic value they can create for a department. Therefore, support for the time-consuming tasks of screening, segmentation, and quantification is likely to penetrate the market most easily. However, one also should bear in mind that value to the patient, the health care system, and society might require additional investment in radiologic diagnostic tools if cost savings can be realized downstream in the care pathway or at a societal level. Therefore, health care systems will face the challenge of incentivizing the development and use of AI tools where they are overall beneficial while avoiding subsidizing tools that increase the cost of unnecessary workup at little or no benefit to patients, as with the first generation of computer-aided detection in breast imaging (11). Solutions that improve the quality of radiologic diagnosis without generating immediate financial benefit are less likely to permeate the mass market rapidly. Nonetheless, their potential to promote health equity through increasing diagnostic quality and consistency in nonsubspecialized settings could profoundly improve the overall performance of a health care system. Therefore, noncommercial stakeholders should pursue research and publication of socially desirable AI solutions.

Tadavarthi et al have highlighted some of the technical, regulatory, economic, and behavioral hurdles that AI-enabled diagnostic support systems need to overcome before widespread clinical adoption. However, we don't need to look far for AI tools that could find a quicker entry into radiology departments. At CT, patient positioning and examination planning contribute significantly to the total examination time; suboptimal imaging is a source of diagnostic errors and repeat examinations in any modality. Similarly, image reconstruction could see rapid adoption of AI algorithms to reduce the scan time, reduce contrast material doses, and increase image quality at MRI or for dose reduction at CT.

AI has reached health care, and radiology in particular, and it is here to stay. Careful evaluation and adoption of AI-based tools will allow radiologists to pioneer the transition toward AI-enabled, patient-centric health care delivery. In collaboration, radiology researchers, health care providers, industry partners, and policymakers have the potential to realize the promise of AI to provide equal access to high-quality care, overcome the challenge of ongoing performance monitoring, and achieve the development of socially beneficial AI solutions.

Disclosures of Conflicts of Interest: E.S. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: author is a research consultant for Amazon; disclosed money paid to author from GSK for lectures and from Siemens for travel/accommodations/meeting expenses unrelated to activities listed; author is co-founder and shareholder of Lucida Medical, but had no earnings. Other relationships: disclosed no relevant relationships. **S.U.** disclosed no relevant relationships.

References

1. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;84(4):465–470.
2. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E. FRAX and the assessment of fracture probability in men and women from the UK. Osteoporos Int 2008;19(4):385–397.
3. Thomassin-Naggara I, Aubert E, Rockall A, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. Radiology 2013;267(2):432–443.
4. Schläpfer J, Wellens HJ. Computer-Interpreted Electrocardiograms: Benefits and Limitations. J Am Coll Cardiol 2017;70(9):1183–1192.

5. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. J Natl Cancer Inst 2019;111(9):916–922.
6. Reyes M, Meier R, Pereira S, et al. On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. Radiol Artif Intell 2020;2(3):e190043.
7. Topol E, Harvey H, Manero E, et al. The Topol Review: Preparing the healthcare workforce to deliver the digital future. Leeds, England: Health Education England, 2019.
8. Tadavarthi Y, Vey B, Krupinski E, et al. The State of Radiology AI – Considerations for Purchase Decisions and Current Market Offerings. Radiol Artif Intell 2020;2(6):e200004.
9. Joshi I, Cushnan D. A Buyer's Guide to AI in Health and Care. https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/. Published 2020. Accessed October 20, 2020.
10. Lee CS, Lee AY. Clinical applications of continual learning machine learning. Lancet Digit Health 2020;2(6):E279–E281.
11. Lehman CD, Wellman RD, Buist DSM, et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Intern Med 2015;175(11):1828–1837.