A survey of ASER members on artificial intelligence in emergency radiology: trends, perceptions, and expectations

Anjali Agrawal[1], **Garvit D. Khatri**[2], **Bharti Khurana**[3], **Aaron D. Sodickson**[3], **Yuanyuan Liang**[4], **David Dreizin**[5]

[1]New Delhi operations, Teleradiology Solutions, Delhi, India

[2]Nuclear Medicine, Department of Radiology, University of Washington School of Medicine, Seattle, WA, USA

[3]Emergency Radiology, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[4]Epidemiology & Public Health, University of Maryland School of Medicine, Baltimore, MD, USA

[5]Trauma and Emergency Radiology, Department of Diagnostic Radiology and Nuclear Medicine, R Adams Cowley Shock Trauma Center, University of Maryland School of Medicine, Baltimore, MD, USA

Abstract

Purpose—There is a growing body of diagnostic performance studies for emergency radiology-related artificial intelligence/machine learning (AI/ML) tools; however, little is known about user preferences, concerns, experiences, expectations, and the degree of penetration of AI tools in emergency radiology. Our aim is to conduct a survey of the current trends, perceptions, and expectations regarding AI among American Society of Emergency Radiology (ASER) members.

Methods—An anonymous and voluntary online survey questionnaire was e-mailed to all ASER members, followed by two reminder e-mails. A descriptive analysis of the data was conducted, and results summarized.

Results—A total of 113 members responded (response rate 12%). The majority were attending radiologists (90%) with greater than 10 years' experience (80%) and from an academic practice (65%). Most (55%) reported use of commercial AI CAD tools in their practice. Workflow prioritization based on pathology detection, injury or disease severity grading and classification, quantitative visualization, and auto-population of structured reports were identified as high-value tasks. Respondents overwhelmingly indicated a need for explainable and verifiable tools (87%) and the need for transparency in the development process (80%). Most respondents did not feel that AI would reduce the need for emergency radiologists in the next two decades (72%) or diminish interest in fellowship programs (58%). Negative perceptions pertained to potential for

---

David Dreizin, daviddreizin@gmail.com.

automation bias (23%), over-diagnosis (16%), poor generalizability (15%), negative impact on training (11%), and impediments to workflow (10%).

Conclusion—ASER member respondents are in general optimistic about the impact of AI in the practice of emergency radiology and its impact on the popularity of emergency radiology as a subspecialty. The majority expect to see transparent and explainable AI models with the radiologist as the decision-maker.

Keywords

Radiology; Imaging; Emergency; Trauma; Emergency radiology; Artificial intelligence; Machine learning; Survey; Computer-aided detection

Introduction

Over the past decade, the development of graphics processing units with rapid parallel computing architectures ushered in an era of scalable multi-layered artificial neural network-based representation learning methods (i.e., deep learning) for computer vision tasks [1, 2]. As a digitized, data-driven field, radiology has been well-positioned toward early adoption of new information technologies [3, 4], and the majority of artificial intelligence and machine learning (AI/ML) software as medical device (SaMD) products are in the radiology domain [4].

Emergency radiology faces a number of unique practice challenges. Services are carried out within a high-stakes time- and safety–critical environment involving ill or injured patients that require expeditious and accurate diagnosis [5, 6]. Reading room distractions are frequent, and off-hours work is associated with performance degradation related to circadian rhythm disruption [5-8]. The availability of a 24-h-a-day workforce with emergency radiology expertise is highly variable across institutions and practices [9-11]. Performance and workflow improvements, and the potential for improved patient outcomes are strong incentives for development of computer-aided detection and diagnosis (CAD) tools in this setting [12, 13].

FDA-approved commercialized products currently include a variety of use cases, primarily for *detection* (CADe) and triage or workflow prioritization (CADt). These include tools for stroke [14], pulmonary embolus [15], intracranial hemorrhage [16], acute pathology on chest radiographs [17], and fractures on musculoskeletal plain radiographs [18]. Additionally, the FDA recognizes tools for diagnosis, risk stratification and prognostication (CADx), and image processing and quantitative visualization (IPQ) [4], which may have a role in augmenting emergency radiology interpretation in the future.

Several studies have evaluated the performance characteristics of commercial CAD tools with respect to diagnostic accuracy [19-21] and effects on turnaround times [16, 22]. Additionally, surveys have been published examining radiologist and radiology trainee views on AI/ML [23-25]. However, little is known about emergency radiologist engagement with AI CAD tools. Furthermore, little is known regarding the degree of penetration of these tools into emergency radiology practice. While members of the emergency radiology community

are aware of the growing use of AI CAD tools through the published literature, society meetings, and social media [26-31], as end-users, and in some cases as developers, a canvas of ASER members may provide valuable insights for AI governance at the institutional and society level, facilitate adoption into emergency radiology workflows, and help with priority setting for development of trustworthy AI systems that maximize clinical impact and adequately meet the needs of emergency radiology end-users.

With the above goals in mind, the American Society of Emergency Radiology AI/ML Expert Panel convened a working group to conduct a survey of its members to better understand emergency radiologists' perceptions and expectations and explore current practice patterns and emerging trends. Dimensions explored included implementation and governance, trust and user acceptance, overall value, unmet needs, and implications for the future of the subspecialty.

Methods

The ASER AI/ML expert panel cross-sectional survey working group designed and conducted an anonymous and voluntary online survey of the ASER membership. The work was determined to be IRB exempt by the primary site's human research protections office.

The survey was intended to gather cross-sectional descriptive information on (a) the practice setting and experience level of respondents; (b) clinical needs, as determined by the value placed on tools performing a variety of clinical tasks; (c) current trends with respect to implementation and governance; (d) dimensions of trust and user acceptance; and (e) expectations and apprehensions for the future of AI in emergency radiology.

Five working group members (including three academic emergency radiologists, one private practice emergency radiologist, and one fellow) reviewed literature on AI in radiology with special focus on applications in emergency radiology and the perceptions regarding AI among the radiology community at large. The group formulated, discussed, commented on, and edited questions in an iterative process to generate a Web-based questionnaire formulated around the aforementioned themes. The survey was created using jotform.com. The final version of the survey consisted of 23 questions of categorical, multiple-choice, 9-point UCLA/RAND-type Likert scale, yes/no, and narrative type with free text for comments. It was designed to be completed in under 15 min.

An initial series of three background questions were used to determine respondents' practice settings (academic, community, teleradiology, or mixed); whether respondents were attendings, fellows, or residents; and years of experience in radiology including residency and fellowship.

Four questions interrogated the current trends of AI tool implementation and governance in practice. Specifically, we asked whether respondents used commercial AI tools at this time; whether the respondents' practices employ processes for local validation or revalidation of deployed tools; whether institutional end-users included radiologists or clinicians, or both; whether the use of AI CAD tools is perceived to have improved the quality of care at

respondents' institutions or practices, and if so, a free-text box was available to provide an optional explanation.

A set of needs-assessment Likert-scale questions was designed to determine the level of impact that respondents expect tools to have on their radiology practice in the future based on the clinical task, including detection and workflow prioritization, disease or injury severity grading, quantitative visualization (e.g., automated measurement of pathology denoted by contours, masks, or electronic caliper measurements), risk stratification/ prognostication, and structured reporting. An open-ended question with free-text response asked respondents to list up to 3 pathologies they would find helpful in the emergency radiology setting. These were bucketed using a free-form approach with major categories determined as responses were processed.

Several position papers on foundational and clinical-translational research have emphasized the importance of AI/ML algorithm trustworthiness both in terms of transparency of the output, and transparency in methodology [32-34]. Several questions were formulated to study the opinion of respondents with respect to system benevolence and trust. These included a Likert question of the level of importance placed on explainable and verifiable (as opposed to black box) results; a Likert question assessing the importance of transparency in ground truth annotation; a Likert question assessing the degree of concern regarding automation bias; and a multiple-choice question asking how often respondents disagree with AI results. Another question probed apprehensions/concerns regarding AI, with a check-box list and free-text option. Choices included concerns about the following: overdiagnosis, workflow, institutional resistance to change, lack of knowledge, cost, ethics, negative impact on training, workflows that bypass radiology, and poor algorithm generalizability to local populations.

Radiologist expectations regarding the future impact of AI/ML on the field were queried with a multiple-choice question regarding whether AI tools will have positive, negative, or no impact on radiologist job satisfaction; a Likert question on the likelihood that AI/ML will reduce the need for 24/7 coverage in the next 20 years; and whether the emergence of AI tools in emergency radiology will impact the interest of residents in pursuing emergency radiology fellowships.

One question pertaining to the importance of human factors engineering in AI research and development was noted to have a design flaw during administration of the survey, preventing analysis of the results, and this is not discussed further.

The Web link was distributed amongst all ASER members via email with a cover letter (Appendix). The initial email was sent on June 3, 2022, with two reminders at 2–4-week intervals, and closing date of August 10, 2022. Another reminder was not deemed necessary as the number of respondents exceeded 100 after the second reminder, the minimum number agreed upon by consensus for a descriptive survey [15].

Median values with interquartile range (IQR) were used to summarize response trends. A descriptive analysis was performed for free-text questions. Question responses were displayed visually where applicable using pie charts, with categories pre-determined from

RAND/UCLA grading scales (e.g., 1–3, not useful; 4–6, uncertain; 7–9, useful). Responses for all categorical and Likert questions were compared by practice type and years of practice subgroups. Comparisons were performed using Fisher exact and Chisquared tests for questions with categorical responses, and the Mann–Whitney U-test and post hoc Kruskal–Wallis H test for questions with Likert scales. Bonferroni adjustment was performed for multiple comparisons. A p-value of < 0.05 was considered significant.

Results

A total of 113 members responded to the survey questionnaire from a total of 955 ASER members (including active, active military, associate, emeritus, fellow, and member in training) for a response rate of 12%. Responses for all yes/no, categorical, multiple choice, and Likert questions are provided in Table 1.

Respondent characteristics (Table 1, questions 1–3)

The majority of respondents worked in an academic practice setting (65%, 74/113), followed by community (19%, 21/113), and teleradiology (7%, 8/113). Hybrid practices, including a mix of community and academic hospitals, accounted for the remaining 9% (10/113).

An overwhelming majority of respondents were attendings (90%, 101/112), followed by fellows (4%, 5/112) and residents (3%, 3/112). The remainder were in private practice as radiologists or partners (3%, 3/112). In total, 45% of radiologists had over 20 years of experience (including years spent in residency and fellowship) (51/112), 35% had 10–20 years of experience (39/112), 18 had 5–10 years of experience (16%), and 4% had spent 1–5 years in radiology practice (4/112).

Implementation and governance (Table 1, questions 4–8)

A slight majority (56%, 63/112) report already using commercial AI tools in their practices. However, only 33% (29/89) of question respondents report that their practices have streamlined AI governance processes in place for ongoing validation or revalidation of implemented tools. While the majority of respondents (66%, 49/74) reported exclusively radiologists as the primary end-users of AI tools at their institution, a third (34%, 25/74) reported both radiologists and clinicians as end-users.

Approximately two-thirds of respondents (64%, 42/66) felt that AI tools have improved the quality of care at their institution. Of 42 respondents that provided a reason for improved care, 71% (30/42) noted improved triage turnaround times, and 57% (24/42) indicated second reader capability. More respondents answered, "who are the primary end-users" (question 6, $n = 74$) and "have tools improved quality of care" (question 7, $n = 66$) than the number who reported using commercial AI tools (question 4, $n = 63$). This may be reconciled by the possibility that some respondents are anticipating installation or that some are trialing non-commercial tools, as well as possible under-reporting by respondents of commercial AI use.

Needs assessment (Table 1, questions 9–14)

Median Likert scores (1–3, no impact; 4–6, some impact; 7–9, high impact) for the perceived future impact of AI CAD tools by task were 7 [IQR = 5, 9] for workflow prioritization and detection, 7 [5, 8] for quantitative visualization, 7 [5, 8] for injury grading and classification, and 7 [5, 9] for AI tools that auto-populate structured reports, but only 5 [4, 7] for tools that provide prognostic information. Analyzing this data categorically, the percentage of respondents that felt that a given tool would have high impact included 61% for detection and workflow prioritization, 58% for quantitative visualization, 60% for disease or injury severity grading or classification, and 61% for auto-population of structured reports, but only 30% for prognostic tools.

For our free-text question soliciting up to 3 pathologies that would be helpful in the emergency setting (see Table 1, question 14), the five most commonly listed major categories of pathology for which AI tools could be helpful, included, in order of most to least common, fractures (47 mentions), pulmonary embolus (39 mentions), features of ischemic stroke (37 mentions), intracranial hemorrhage (31 mentions), and intracavitary torso hemorrhage-related features (21 mentions). Of all fracture-related tools, AI tools for rib fracture detection and numbering were felt to be the most useful (20 of 46 mentions).

System benevolence and trust (Table 1, questions 15–19)

An overwhelming majority (87%, 98/113) gave high priority (Likert scores of 7–9) to AI tools with interpretable and verifiable results that can be rejected when perceived to be incorrect by the end-user (median = 9; IQR [8, 9]). A similarly high percentage (80%, 86/108) of respondents indicated that since AI tools are trained using expert annotation, and ground-truth agreement between experts can vary considerably by task, it is very important to know the level of expert annotation agreement (median = 8; IQR [7, 9]). In total, 11% (12/108) were uncertain on this matter.

While 39% (40/101) have no concerns of AI tools resulting in a biased image interpretation, 28% (28/101) had a high level of concern regarding the possibility of automation bias being introduced by AI tools during interpretation (median = 5, IQR [3, 7]).

Regarding frequency of disagreement with AI tool results, respondents most commonly reported disagreeing with AI output in 5–10% of studies (37%, 24/65), followed by 10–20% of studies (35%, 23/65). More extreme experiences were less common, but similar at both ends of the spectrum, with 12% (8/65) reporting disagreeing with diagnostic AI tool results in fewer than 5% of studies, and 16% (10/65) reporting disagreeing in greater than 20%.

The most common specific concerns with respect to AI tools in the emergency radiology setting included overdiagnosis (61%, 63/103), non-generalizability of published performance (57%, 59/103), negative impact on training (44%, 45/103), impediments to the workflow (41%, 42/103), and insufficient evidence to support use (40%, 41/103).

Expectations (Table 1, questions 20–22)

While 72% of respondents (78/109) felt that AI tools will increase radiologist job satisfaction, 10% (11/109) felt that job satisfaction would be negatively impacted.

Approximately three-quarters (73%, 82/113) felt that the likelihood that AI/ML would reduce the need for 24/7 emergency radiology coverage in the next two decades was low, and the majority (60%, 65/108) felt that AI would have no impact on interest in pursuing emergency radiology fellowship among radiology residents. Interestingly, a third of the respondents (33%, 35/108) felt that AI would lead to an increased interest in the subspecialty among trainees.

After Bonferroni correction, subgroup analyses of responses by practice type, and years of practice found a statistically significant difference only with respect to the need for transparency in reporting of reader agreement, with higher priority given to this aspect of methodological transparency by academic radiologists compared with community radiologists ($p = 0.024$).

Discussion

This survey of ASER members was conducted to gain insights about current AI/ML trends, perceptions, and expectations in emergency radiology. Our results reveal that, in line with responses to prior radiology surveys [23-25], most respondents that currently use AI tools are positive about their potential patient care value. However, respondents have concerns about the capability and benevolence of tools, particularly with respect to overdiagnosis and generalizability. Guarded optimism regarding AI as a value-adding technology in radiology comes at a time when there are still few but rapidly increasing numbers of FDA-approved AI tools in the emergency radiology domain that have efficacy as second-reader tools [19-21], and triage and notification tools that reduce turnaround times, and potentially patient length of stay [16].

In questions pertaining to trust and system benevolence, respondents indicated overwhelmingly that AI tools must be explainable and verifiable. Extremes on either end of reported levels of disagreement (less than 5% or over 20%) with AI output were uncommon in this survey, with the majority of respondents reporting disagreement rates between 5 and 20%, suggesting a perceived high level of system capability for commercial tools currently in use. Most commercial tools employ activation maps, box detections, segmentations, or annotations for explainability [1]. In the near-term, and perhaps indefinitely, it is critical for humans to remain in the drivers' seat when performing AI-assisted reads. This is best ensured through interpretable AI output that can be verified or rejected, and potentially with greater levels of human-in-the-loop interaction [33-36].

Needs assessment questions reveal that approximately two-thirds of respondents place a high value on future tools that perform triage and early notification, quantitative visualization, grading and classification, and auto-population of structured reports. In free-text responses, pathologies where AI is most likely to be helpful include fractures, pulmonary embolus, ischemic stroke, intracranial hemorrhage, and torso hemorrhage. Currently available commercial tools for the most part perform simple detection tasks, with detection/second reader (CADe) and triage/early notification (CADt) intended uses. Commercialized tools with regulatory approval cover four of the five major pathology categories, including fractures, pulmonary embolus, ischemic stroke, and intracranial hemorrhage [14, 15, 37-42].

Grading and classification of pathology is a complex task involving multiple diagnostic steps. Tools for this purpose are currently rare or appear to be in early stages of the research and development pipeline. Interpretability is also more difficult to satisfy for such multi-stage methods [13, 35, 43, 44]. There are also few tools that meet the need for quantification of disease or injury. Such tools are considered beneficial from a precision medicine standpoint [2]. To our knowledge, there are currently no commercial products that meet the need for torso hemorrhage-related pathologies. As the torso occupies a large volume, and targets are often small but highly variable in volume and appearance, multi-scale algorithms for complex torso pathology have been latecomers in the era of deep learning [45-53].

Respondents were less enthusiastic for algorithms that prognosticate outcomes. Speculative reasons for this may include (i) scope of practice for this fast-paced and high-volume subspecialty that mainly focuses on timely diagnosis on the front end, with decision-making based on risk or prognosis left to treating members of the care team; and (ii) lack of trust, as prognostication tools often have a black box problem and potential ethical concerns [54].

Transparency throughout the research and development process including data curation is also found to be a top priority for our respondents. Concerns with research transparency, and other specific apprehensions such as overdiagnosis, negative impact on training, and ethical concerns, which include the possibility of selective misdiagnosis in underserved patient populations, are consistent with concerns previously raised in the medical and radiology literature [34, 36] and should be taken seriously by solution developers [13, 55-57].

Slightly more than half of respondents currently use commercial AI tools and few reported streamlined processes for local validation/re-validation of implemented tools. Adoption of AI tools is known to be hampered by the lack of business incentive in a fee-for-service environment and lack of outcome data for reimbursement. Patient outcome data is required for reimbursement through the Centers for Medicare and Medicaid Services New Technology Add-on Payment (CMS NTAP). High costs and lack of reimbursement may make it difficult for those involved in AI governance to justify the expense of AI CAD tools to hospital administrators [58-60]. Approximately one-third of respondents indicate both radiologist and clinician end-users within their institutions or practices. Those seeking administrative support for the acquisition of AI products may potentially consider buy-in from clinical stakeholders. This will likely become more important as value-based payment models evolve.

Despite the aforementioned specific concerns, respondents in our survey largely felt that AI is unlikely to displace the round-the-clock emergency radiology coverage model, and unlikely to dissuade trainees from pursuing a career in the subspecialty. Some even expect emergency radiology to become more popular as a subspecialty on the cutting edge of AI. This is contrary to the reported skepticism towards AI in radiology by medical students who perceive it as a potential threat to diagnostic radiologists and one of the reasons for not pursuing radiology as a specialty [25, 61]. This may be explained in part by the characteristics of the respondents, as most were attendings with more than 10 years of experience, and there were few trainee respondents.

Bias was not a major concern for most respondents; however, automation and complacency bias are acknowledged pitfalls of AI implementation in the radiology and medical AI/ML literature [62, 63]. Radiologists, being only human, may have a blind spot for their own biases, and avoiding bias-related error has been an area of interest in the radiology literature since well before the arrival of AI/ML [64, 65]. Greater awareness of these problems is likely warranted.

Our survey was limited by the low response rate. While anonymous and voluntary, this survey interrogated a select group of emergency radiologists, with responses coming disproportionately from practitioners in the academic setting. In preserving anonymity, we were not able to assess entries by respondents' institutions, which could potentially skew results towards over-representation by the perspectives and preferences of larger or more AI-engaged institutions. Furthermore, respondents, who were disproportionately attendings and had over 10 years of experience, perhaps share a common outlook towards radiology and AI/ML that informs their level of acceptance of artificial intelligence technologies and concerns. While demographic differences between respondents and non-respondents were not explored, it is likely that only a subset of ASER members interested in AI were motivated to respond to our survey, leading to a selection bias which limits generalization of results to the emergency radiology community at large.

Conclusion

Just over half of respondents among the ASER membership currently use commercial AI tools in their practice. Two-thirds of respondents who currently use AI tools feel that they improve quality of care, and most find themselves disagreeing with AI predictions in 5–20% of studies. Concerns and apprehensions pertaining to overdiagnosis and generalization to their local patient populations are shared by over half of end-users. The majority of respondents expect to see transparent and explainable AI tools with the onus of the final decision with the radiologist.

Unmet needs identified included tools that perform grading or classification, quantitative visualization tasks, and tools that auto-populate structured reports. Torso hemorrhage-related tools were a commonly listed pathology for which AI tools could be helpful, and we are presently not aware of commercial FDA-approved tools in this area.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

References

1. Fujita H (2020) AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiol Phys Technol 13(1):6–19 [PubMed: 31898014]

2. Zhou SK, Greenspan H, Davatzikos C, Duncan JS, Van Ginneken B, Madabhushi A, Prince JL, Rueckert D, Summers RM (2021) A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. Proc IEEE 109(5):820–838

3. West E, Mutasa S, Zhu Z, Ha R (2019) Global trend in artificial intelligence-based publications in radiology from 2000 to 2018. Am J Roentgenol 213(6):1204–1206 [PubMed: 31414886]

4. Ebrahimian S, Kalra MK, Agarwal S, Bizzo BC, Elkholy M, Wald C, Allen B, Dreyer KJ (2022) FDA-regulated AI algorithms: trends, strengths, and gaps of validation studies. Acad Radiol 29(4):559–566 [PubMed: 34969610]

5. Banaste N, Caurier B, Bratan F, Bergerot J-F, Thomson V, Millet I (2018) Whole-body CT in patients with multiple traumas: factors leading to missed injury. Radiology 289(2):374–383 [PubMed: 30084754]

6. Hanna TN, Zygmont ME, Peterson R, Theriot D, Shekhani H, Johnson J-O, Krupinski EA (2018) The effects of fatigue from overnight shifts on radiology search patterns and diagnostic performance. J Am Coll Radiol 15(12):1709–1716 [PubMed: 29366599]

7. Bruno MA (2020) Radiology errors across the diurnal cycle. Radiology 297(2):380–381. 10.1148/radiol.2020202902 [PubMed: 32813599]

8. Glover M IV, Almeida RR, Schaefer PW, Lev MH, Mehan WA Jr (2017) Quantifying the impact of noninterpretive tasks on radiology report turn-around times. J Am Coll Radiol 14(11):1498–1503 [PubMed: 28916177]

9. Chong ST, Robinson JD, Davis MA, Bruno MA, Roberge EA, Reddy S, Pyatt RS Jr, Friedberg EB (2019) Emergency radiology: current challenges and preparing for continued growth. J Am Coll Radiol 16(10):1447–1455 [PubMed: 31092353]

10. Hanna TN, Shekhani H, Lamoureux C, Mar H, Nicola R, Sliker C, Johnson J-O (2017) Emergency radiology practice patterns: shifts, schedules, and job satisfaction. J Am Coll Radiol 14(3):345–352 [PubMed: 27927590]

11. Kalyanpur A, Weinberg J, Neklesa V, Brink JA, Forman HP (2003) Emergency radiology coverage: technical and clinical feasibility of an international teleradiology model. Emerg Radiol 10(3):115–118 [PubMed: 15290497]

12. Kalyanpur A (2020) Teleradiology and artificial intelligence–birds of the same feather. Acad Radiol 27(1):123–126 [PubMed: 31147240]

13. Agrawal A (2022) Emergency teleradiology-past, present, and is there a future? Front Radiol 2:866643

14. Soun J, Chow D, Nagamine M, Takhtawala R, Filippi C, Yu W, Chang P (2021) Artificial intelligence and acute stroke imaging. Am J Neuroradiol 42(1):2–11 [PubMed: 33243898]

15. Soffer S, Klang E, Shimon O, Barash Y, Cahan N, Greenspana H, Konen E (2021) Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis. Sci Rep 11(1):1–8 [PubMed: 33414495]

16. Davis MA, Rao B, Cedeno PA, Saha A, Zohrabian VM (2022) Machine learning and improved quality metrics in acute intracranial hemorrhage by noncontrastcomputed tomography. Curr Probl Diagn Radiol 51(4):556–561. 10.1067/j.cpradiol.2020.10.007 [PubMed: 33243455]

17. Gipson J, Tang V, Seah J, Kavnoudias H, Zia A, Lee R, Mitra B, Clements W (2022) Diagnostic accuracy of a commercially available deep-learning algorithm in supine chest radiographs following trauma. Br J Radiol 95:20210979 [PubMed: 35271382]

18. Guermazi A, Tannoury C, Kompel AJ, Murakami AM, Ducarouge A, Gillibert A, Li X, Tournier A, Lahoud Y, Jarraya M (2022) Improving radiographic fracture recognition performance and efficiency using artificial intelligence. Radiology 302(3):627–636 [PubMed: 34931859]

19. Seah JC, Tang CH, Buchlak QD, Holt XG, Wardman JB, Aimoldin A, Esmaili N, Ahmad H, Pham H, Lambert JF (2021) Effect of a comprehensive deep-learning model on the accuracy of chest X-ray interpretation by radiologists: a retrospective, multireader multicase study. Lancet Digit Health 3(8):e496–e506 [PubMed: 34219054]

20. Kau T, Ziurlys M, Taschwer M, Kloss-Brandstätter A, Grabner G, Deutschmann H (2022) FDA-approved deep learning software application versus radiologists with different levels of expertise:

detection of intracranial hemorrhage in a retrospective singlecenter study. Neuroradiology 64(5):981–990 [PubMed: 34988593]

21. Duron L, Ducarouge A, Gillibert A, Lainé J, Allouche C, Cherel N, Zhang Z, Nitche N, Lacave E, Pourchot A (2021) Assessment of an AI aid in detection of adult appendicular skeletal fractures by emergency physicians and radiologists: a multicenter crosssectional diagnostic study. Radiology 300(1):120–129 [PubMed: 33944629]

22. Wismüller A, Stockmaster L (2020) A prospective randomized clinical trial for measuring radiology study reporting time on Artificial Intelligence-based detection of intracranial hemorrhage in emergent care head CT. In: Proc. SPIE 11317, Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging, p 113170M. 10.1117/12.2552400

23. Huisman M, Ranschaert E, Parker W, Mastrodicasa D, Koci M, Pinto de Santos D, Coppola F, Morozov S, Zins M, Bohyn C (2021) An international survey on AI in radiology in 1,041 radiologists and radiology residents part 1: fear of replacement, knowledge, and attitude. Eur Radiol 31(9):7058–7066 [PubMed: 33744991]

24. Huisman M, Ranschaert E, Parker W, Mastrodicasa D, Koci M, Pinto de Santos D, Coppola F, Morozov S, Zins M, Bohyn C (2021) An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education. Eur Radiol 31(11):8797–8806 [PubMed: 33974148]

25. van Hoek J, Huber A, Leichtle A, Härmä K, Hilt D, von Tengg-Kobligk H, Heverhagen J, Poellinger A (2019) A survey on the future of radiology among radiologists, medical students and surgeons: students and surgeons tend to be more skeptical about artificial intelligence and radiologists may fear that other disciplines take over. Eur J Radiol 121:108742 [PubMed: 31734640]

26. Jalal S, Parker W, Ferguson D, Nicolaou S (2021) Exploring the role of artificial intelligence in an emergency and trauma radiology department. Can Assoc Radiol J 72(1):167–174 [PubMed: 32309989]

27. Moulik SK, Kotter N, Fishman EK (2020) Applications of artificial intelligence in the emergency department. Emerg Radiol 27:355–358. 10.1007/s10140-020-01794-1 [PubMed: 32643067]

28. Jacques T, Fournier L, Zins M, Adamsbaum C, Chaumoitre K, Feydy A, Millet I, Montaudon M, Beregi J-P, Bartoli J-M (2021) Proposals for the use of artificial intelligence in emergency radiology. Diagn Interv Imaging 102(2):63–68 [PubMed: 33279461]

29. Lakhani P, Prater AB, Hutson RK, Andriole KP, Dreyer KJ, Morey J, Prevedello LM, Clark TJ, Geis JR, Itri JN (2018) Machine learning in radiology: applications beyond image interpretation. J Am Coll Radiol 15(2):350–359 [PubMed: 29158061]

30. Noguerol TM, Paulano-Godino F, Martín-Valdivia MT, Menias CO, Luna A (2019) Strengths, weaknesses, opportunities, and threats analysis of artificial intelligence and machine learning applications in radiology. J Am Coll Radiol 16(9):1239–1247 [PubMed: 31492401]

31. Goldberg JE, Rosenkrantz AB (2019) Artificial intelligence and radiology: a social media perspective. Curr Probl Diagn Radiol 48(4):308–311 [PubMed: 30143386]

32. Langlotz CP, Allen B, Erickson BJ, Kalpathy-Cramer J, Bigelow K, Cook TS, Flanders AE, Lungren MP, Mendelson DS, Rudie JD (2019) A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology 291(3):781 [PubMed: 30990384]

33. Allen B Jr, Seltzer SE, Langlotz CP, Dreyer KP, Summers RM, Petrick N, Marinac-Dabic D, Cruz M, Alkasab TK, Hanisch RJ (2019) A road map for translational research on artificial intelligence in medical imaging: from the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop. J Am Coll Radiol 16(9):1179–1189 [PubMed: 31151893]

34. Park SH, Han K (2018) Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Radiology 286(3):800–809 [PubMed: 29309734]

35. Chen H, Gomez C, Huang C-M, Unberath M (2022) Explainable medical imaging AI needs human-centered design: guidelines and evidence from a systematic review. npj Digit Med 5(1):1–15 [PubMed: 35013539]

36. Bluemke DA, Moy L, Bredella MA, Ertl-Wagner BB, Fowler KJ, Goh VJ, Halpern EF, Hess CP, Schiebler ML, Weiss CR (2020) assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers-from the radiology editorialboard. Radiology 294(3):487–489. 10.1148/radiol.2019192515 [PubMed: 31891322]

37. Jones RM, Sharma A, Hotchkiss R, Sperling JW, Hamburger J, Ledig C, O'Toole R, Gardner M, Venkatesh S, Roberts MM (2020) Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. NPJ Digit Med 3(1):1–6 [PubMed: 31934645]

38. Dupuis M, Delbos L, Veil R, Adamsbaum C (2022) External validation of a commercially available deep learning algorithm for fracture detection in children. Diagn Interv Imaging 103(3):151–159 [PubMed: 34810137]

39. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, Mahajan V, Rao P, Warier P (2018) Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 392(10162):2388–2396 [PubMed: 30318264]

40. Ginat DT (2020) Analysis of head CT scans flagged by deep learning software for acute intracranial hemorrhage. Neuroradiology 62(3):335–340 [PubMed: 31828361]

41. Voter A, Larson M, Garrett J, Yu J-P (2021) Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. Am J Neuroradiol 42(8):1550–1556 [PubMed: 34117018]

42. Voter AF, Meram E, Garrett JW, John-Paul JY (2021) Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of intracranial hemorrhage. J Am Coll Radiol 18(8):1143–1152 [PubMed: 33819478]

43. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R (2020) Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 58:82–115

44. Lavin A, Gilligan-Lee CM, Visnjic A, Ganju S, Newman D, Ganguly S, Lange D, Baydin AG, Sharma A, Gibson A (2022) Technology readiness levels for machine learning systems. Nat Commun 13(1):1–19 [PubMed: 34983933]

45. Lee S, Summers RM (2021) Clinical artificial intelligence applications in radiology: chest and abdomen. Radiol Clin 59(6):987–1002

46. Dreizin D, Zhou Y, Fu S, Wang Y, Li G, Champ K, Siegel E, Wang Z, Chen T, Yuille AL (2020) A Multiscale Deep Learning Method for Quantitative Visualization of Traumatic Hemoperitoneum at CT: Assessment of Feasibility and Comparison with Subjective Categorical Estimation. Radiol Artif Intell 2(6):e190220 [PubMed: 33330848]

47. Dreizin D, Zhou Y, Zhang Y, Tirada N, Yuille AL (2020) Performance of a deep learning algorithm for automated segmentation and quantification of traumatic pelvic hematomas on CT. J Digit Imaging 33(1):243–251 [PubMed: 31172331]

48. Chen H, Unberath M, Dreizin D (2023) Toward automated interpretable AAST grading for blunt splenic injury. Emerg Radiol 30:41–50. 10.1007/s10140-022-02099-1 [PubMed: 36371579]

49. Zhou Y, Dreizin D, Wang Y, Liu F, Shen W, Yuille AL (2021) External attention assisted multi-phase splenic vascular injury segmentation with limited data. IEEE Trans Med Imaging 41(6):1346–1357

50. Dreizin D, Zhou Y, Chen T, Li G, Yuille AL, McLenithan A, Morrison JJ (2020) Deep learning-based quantitative visualization and measurement of extraperitoneal hematoma volumes in patients with pelvic fractures: potential role in personalized forecasting and decision support. J Trauma Acute Care Surg 88(3):425 [PubMed: 32107356]

51. Dreizin D, Chen T, Liang Y, Zhou Y, Paes F, Wang Y, Yuille AL, Roth P, Champ K, Li G (2021) Added value of deep learning-based liver parenchymal CT volumetry for predicting major arterial injury after blunt hepatic trauma: a decision tree analysis. Abdom Radiol 46(6):2556–2566

52. Choi J, Mavrommati K, Li NY, Patil A, Chen K, Hindin DI, Forrester JD (2022) Scalable deep learning algorithm to compute percent pulmonary contusion among patients with rib fractures. J Trauma Acute Care Surg 93(4):461–466 [PubMed: 35319542]

53. Röhrich S, Hofmanninger J, Negrin L, Langs G, Prosch H (2021) Radiomics score predicts acute respiratory distress syndrome based on the initial CT scan after trauma. Eur Radiol 31(8):5443–5453 [PubMed: 33733689]

54. Wang F, Kaushal R, Khullar D (2020) Should health care demand interpretable artificial intelligence or accept "Black Box" Medicine? Ann Intern Med 172(1):59–60. 10.7326/M19-2548 [PubMed: 31842204]

55. Adamson AS, Welch HG (2019) Machine learning and the cancer-diagnosis problem-no gold standard. N Engl J Med 381(24):2285–2287 [PubMed: 31826337]

56. Banerjee M, Chiew D, Patel KT, Johns I, Chappell D, Linton N, Cole GD, Francis DP, Szram J, Ross J (2021) The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. BMC Med Educ 21(1):1–10 [PubMed: 33388043]

57. Seyyed-Kalantari L, Zhang H, McDermott M, Chen IY, Ghassemi M (2021) Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med 27(12):2176–2182 [PubMed: 34893776]

58. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K (2019) The practical implementation of artificial intelligence technologies in medicine. Nat Med 25(1):30–36 [PubMed: 30617336]

59. Daye D, Wiggins WF, Lungren MP, Alkasab T, Kottler N, Allen B, Roth CJ, Bizzo BC, Durniak K, Brink JA (2022) Implementation of clinical artificial intelligence in radiology: who decides and how? Radiology 305(3):555–563 [PubMed: 35916673]

60. Lin M (2022) What's needed to bridge the gap between US FDA Clearance and real-world use of AI algorithms. Acad Radiol 29(4):567–568 [PubMed: 34794879]

61. Bin Dahmash A, Alabdulkareem M, Alfutais A, Kamel AM, Alkholaiwi F, Alshehri S, Al Zahrani Y, Almoaiqel M (2020) Artificial intelligence in radiology: does it impact medical students preference for radiology as their future career? BJR| Open 2:20200037 [PubMed: 33367198]

62. Ellahham S, Ellahham N, Simsekler MCE (2020) Application of artificial intelligence in the health care safety context: opportunities and challenges. Am J Med Qual 35(4):341–348 [PubMed: 31581790]

63. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K (2019) Artificial intelligence, bias and clinical safety. BMJ Qual Saf 28(3):231–237

64. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE (2013) Cognitive and system factors contributing to diagnostic errors in radiology. Am J Roentgenol 201(3):611–617 [PubMed: 23971454]

65. Patlas MN, Katz DS, Scaglione M (2019) Errors in emergency and trauma radiology: Springer

Table 1

Questions and summary of responses

Questions (*n* = number of respondents)	Responses	Number of respondents (*n*, (%))	Median Likert score [IQR]
Respondent characteristics			
1. How would you describe your practice? (*n* = 113)	Academic	74 (65%)	
	Community	21 (19%)	
	Teleradiology	8 (7%)	
	Mixed	10 (9%)	
2. Are you an attending, fellow, or resident? (*n* = 112)	Attending	101 (90%)	
	Fellow	5 (4%)	
	Resident	3 (3%)	
	Other	3 (3%)	
3. How many years have you practiced radiology? (*n* = 112)	> 20 years	51 (45%)	
	> 10–20 years	39 (35%)	
	5–10 years	18 (16%)	
	Less than 5 years	4 (4%)	
Implementation and governance			
4. Do you use commercial AI tools in your practice? (*n* = 112)	Yes	63 (56%)	
	No	49 (44%)	
5. Does your practice have streamlined processes in place to perform ongoing local validation/revalidation of implemented tools? (*n* = 89)	Yes	29 (33%)	
	No	60 (67%)	
6. Who are the primary end-users of AI tools at your institution? (*n* = 74)	Radiologists	49 (66%)	
	Radiologists and clinicians	25 (34%)	
7. Have AI CAD tools in use at your institution improved quality of care? (*n* = 66)	Yes	42 (64%)	
	No	24 (36%)	
8. If yes to above, in what way? (*n* = 42) [†]	Improving triage and turnaround	24 (57%)	
	Providing second reader capability	30 (71%)	
	Other	0 (0%)	
Needs assessment			

Questions (*n* = number of respondents)	Responses	Number of respondents (*n*, (%))	Median Likert score [IQR]
Rate the level of impact the following AI tools could have on your practice in the future			
9. AI CAD tools that help with workflow prioritization based on detected pathology (*n* = 113)	High impact	69 (61%)	7 [5, 9]
	Some impact	31 (27%)	
	No impact	13 (12%)	
10. AI CAD tools that quantify pathology (*n* = 113)	High impact	65 (58%)	7 [5, 8]
	Some impact	32 (28%)	
	No impact	16 (14%)	
11. AI CAD tools that assist in grading injury or disease severity based on established classification systems (*n* = 112)	High impact	67 (60%)	7 [5, 8]
	Some impact	32 (28%)	
	No impact	13 (12%)	
12. AI CAD tools that provide prognostic information such as probability of poor clinical outcome (*n* = 113)	High impact	34 (30%)	5 [4, 7]
	Some impact	52 (46%)	
	No impact	27 (24%)	
13. AI tools that auto-populate structured reports (*n* = 113)	High impact	69 (61%)	7 [5, 9]
	Some impact	28 (25%)	
	No impact	16 (14%)	
14. List up to 3 pathologies for which you believe AI tools will be helpful in the ER	Top 5 major categories (collated)	Number of free-response mentions	
	1. Fractures	47	
	Rib	20	
	General	19	
	Spine	5	
	Pelvis	3	
	2. Pulmonary embolus	39	
	3. Ischemic stroke	37	
	General	32	
	Large vessel occlusion	4	
	Perfusion imaging	1	
	4. Intracranial hemorrhage	31	

Questions (n = number of respondents)	Responses	Number of respondents (n, (%))	Median Likert score [IQR]
	5. Intracavitary torso hemorrhage-related	21	
	Solid organ laceration	8	
	General	3	
	Active extravasation	3	
	Gastrointestinal bleed	3	
	Hemoperitoneum	2	
	Hemothorax	1	
	Aortic injury	1	
System benevolence and trust			
15. How important is it to you that AI tools provide interpretable/verifiable results that can be rejected when perceived to be incorrect by the end-user? (1–3, not important; 4–6, uncertain; 7–9, very important) (n = 113)	Very important	98 (87%)	9 [8, 9]
	Uncertain	10 (9%)	
	Not important	5 (4%)	
16. AI tools are trained using expert annotation, and ground-truth agreement between experts can vary considerably by task. How important is it for you to know the level of expert annotation agreement? (1–3, not important; 4–6, uncertain; 7–9, very important) (n = 108)	Very important	86 (80%)	8 [7, 9]
	Uncertain	12 (11%)	
	Not important	10 (9%)	
17. Do you have any concerns that AI tools bias your interpretation of images? (1–3, no; 4–6, uncertain; 7–9, yes) (n = 101)	Yes	28 (28%)	5 [3, 7]
	No	33 (33%)	
	Uncertain	40 (39%)	
18. How often do you find yourself disagreeing with diagnostic AI tool results? (n = 65)	<5% of studies	8 (12%)	
	5–10% of studies	24 (37%)	
	10–20% of studies	23 (35%)	
	<20% of studies	10 (16%)	
19. Do you have any of the following apprehensions/concerns with respect to AI tools in the ER setting? Select all that apply (n = 103): †	-Overdiagnosis	63 (61%)	
	-Reported performance may not generalize to local performance	59 (57%)	
	-Negatively impacts training	45 (44%)	
	-May slow workflow	42 (41%)	
	-Not enough data in literature to support its use	41 (40%)	

Questions (n = number of respondents)	Responses	Number of respondents (n, (%))	Median Likert score [IQR]
	-AI workflow bypasses radiology	36 (35%)	
	-Waste of money	34 (33%)	
	-Not enough knowledge available	26 (25%)	
	-Ethics concerns	26 (25%)	
	-Institution not capable of change	23 (22%)	
Expectations			
20. What impact, if any, do you expect AI tools will have on radiologist job satisfaction (n = 109)	Increased	78 (72%)	
	Decreased	11 (10%)	
	No impact	20 (18%)	
21. On a scale of 1–9, how likely is AI to reduce the need for 24/7 emergency radiology coverage in the next 20 years? (n = 113)	Likely	9 (8%)	2 [1, 4]
	Uncertain	22 (19%)	
	Unlikely	82 (73%)	
22. Will the emergence of AI tools in ER impact interest in pursuing emergency radiology fellowship among radiology residents? (n = 108)	Decreased interest	8 (7%)	
	Increased interest	35 (33%)	
	No impact	65 (60%)	

[†] Numerator = total number of respondents. More than one entry can be provided per respondent