European Journal of Radiology 191 (2025) 112311

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Review

Recognising errors in AI implementation in radiology: A narrative review

Nikolaos Stogiannos [a,b,c,*,1], Renato Cuocolo [d], Tugba Akinci D'Antonoli [e,f], Daniel Pinto dos Santos [g], Hugh Harvey [h], Merel Huisman [i], Burak Kocak [j], Elmar Kotter [k], Karim Lekadir [l,m], Susan Cheng Shelmerdine [n,o,p], Kicky G van Leeuwen [q,r], Peter van Ooijen [s], Michail E. Klontzas [t,u,v], Christina Malamateniou [a,c,w,x]

[a] *Department of Midwifery & Radiography, C.R.R.A.G. Research Group, School of Health and Medical Sciences, City St George's, University of London, London, UK*
[b] *Magnitiki Tomografia Kerkiras, Corfu, Greece*
[c] *European Federation of Radiographer Societies (EFRS), Cumieira, Portugal*
[d] *Department of Medicine, Surgery and Dentistry, University of Salerno, Baronissi, Italy*
[e] *Department of Diagnostic and Interventional Neuroradiology, University Hospital Basel, Basel, Switzerland*
[f] *Department of Pediatric Radiology, University Children's Hospital Basel, Basel, Switzerland*
[g] *Department of Radiology, University Medical Center Mainz, Mainz, Germany*
[h] *Hardian Health, United Kingdom*
[i] *Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, the Netherlands*
[j] *Department of Radiology, Basaksehir Cam and Sakura City Hospital, Istanbul, Turkey*
[k] *Department of Diagnostic and Interventional Radiology, Medical Center, University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany*
[l] *Department of Mathematics and Computer Science, Artificial Intelligence in Medicine Lab (BCN-AIM), Universitat de Barcelona, Barcelona, Spain*
[m] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*
[n] *Department of Clinical Radiology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, United Kingdom*
[o] *University College London Great Ormond Street Institute of Child Health, London, United Kingdom*
[p] *NIHR Great Ormond Street Hospital Biomedical Research Centre, Bloomsbury, London, United Kingdom*
[q] *Romion Health, Utrecht, the Netherlands*
[r] *Health AI Register, Utrecht, the Netherlands*
[s] *Department of Radiotherapy and Data Science Center in Health, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands*
[t] *Artificial Intelligence and Translational Imaging (ATI) Laboratory, Department of Radiology, Medical School, University of Crete, Voutes, Heraklion, Crete, Greece*
[u] *Division of Radiology, Department for Clinical Science Intervention and Technology (CLINTEC), Karolinska Institutet, Stockholm, Sweden*
[v] *Computational Biomedicine Laboratory, Institute of Computer Science, Foundation for Research and Technology (FORTH), Crete, Greece*
[w] *Department of Neuroimaging, King's College London, London, United Kingdom*
[x] *European Society of Medical Imaging Informatics (EuSoMII), Vienna, Austria*

ARTICLE INFO

Keywords:
Artificial Intelligence
Radiology
Failures
Errors
Implementation

ABSTRACT

The implementation of AI can suffer from a wide variety of failures. These failures can impact the performance of AI algorithms, impede the adoption of AI solutions in clinical practice, lead to workflow delays, or create unnecessary costs. This narrative review aims to comprehensively discuss different reasons for AI failures in Radiology through the analysis of published evidence across three main components of AI implementation: (i) the AI models throughout their lifecycle, (ii) the technical infrastructure, including the hardware and software needed to develop and deploy AI models and (iii) the human factors involved. Ultimately, based on the identified errors, this report aims to propose solutions to optimise the use and adoption of AI in radiology.

1. Introduction

The implementation of AI-enabled solutions, either in synergy with human users or as a standalone tool, have exhibited wide performance variability in different clinical contexts and scenarios [1,2]. This highlights the importance of careful planning for the integration of AI in clinical workflows, of continued post-market surveillance, and of simultaneously investing on AI literacy building within healthcare

* Corresponding author.
E-mail address: nikos.stogiannos@citystgeorges.ac.uk (N. Stogiannos).
1 Postal address: Northampton square, London EC1V 0HB.

https://doi.org/10.1016/j.ejrad.2025.112311
Received 6 June 2025; Accepted 8 July 2025
Available online 10 July 2025

organizations [3,4]. Different challenges exist when implementing AI solutions in clinical radiology practice. In many cases, AI models present flaws when faced with unseen, real-world data and complex clinical scenarios [5]. In other cases, failures can occur due to trust miscalibration, resulting from a combination of overreliance on technology and/or lack of user experience, which may lead to wrong diagnoses, poor patient outcomes or, even, patient harm. AI erroneous result may also have the opposite effect, due to algorithmic aversion, which may lead to underutilization of already available reliable resources or create unnecessary costs [6,7]. Academic dissemination is prone to publication bias, thus shunning any negative results from research projects, further contributing to the "file drawer problem" commonly known in science [8,9]. Non-reporting of negative findings on AI implementation [10] inevitably results in unnecessary repeats of experiments or of studies that have failed before, and leads to suboptimal use of resources (time, funding, and human effort). Another form of implementation failure occurs when the AI tool is not aligned with the clinical care pathways of the institution, or when integration into existing workflows has not been adequately considered. Finally, simple shortcomings in infrastructure, like connectivity or compatibility between hardware or software may mean that AI implementation might stall despite a robust and fully tested AI product, a solid business case, an expert team. To understand AI implementation failures, it may be more effective to reverse-engineer success by identifying essential components for seamless AI integration and adoption and tracing back where errors occur. These essential components include:

a) **The AI models employed** across their lifecycle, from inception to decommissioning (see also Fig. 1).

b) **The technical infrastructure**. This is a sparsely studied area which can relate to a range of issues such as connectivity, interoperability, compatibility, etc; therefore, it requires multidisciplinary collaboration and more attention. Frameworks such as the machine learning operations (MLOps) can be used to assist in standardised and seamless deployment practices and prevent AI Implementation mistakes related to infrastructure [11].

c) **The human factors involved**. Human factors relate to a complex system of factors affecting human performance, communication and collaboration, such as physical, cognitive, perceptual, emotional, procedural, and sociocultural [12].

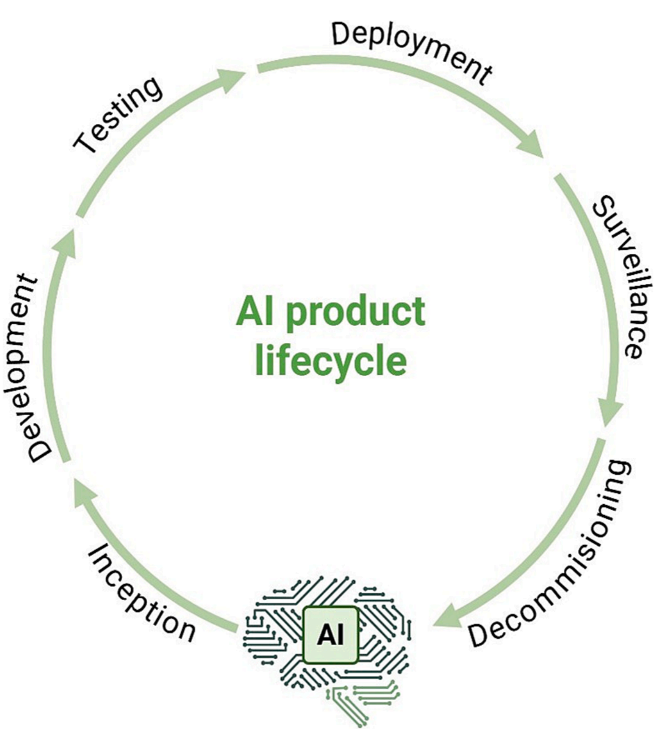

Fig. 1. AI product lifecycle (Created with Biorender).

This paper aims to explore the reasons of AI implementation errors in Radiology and to discuss potential remedial actions to counteract these, by addressing all three essential components and their constituents, as described above. A diverse group of experts in medical imaging met online to agree the aim of the paper and discuss the definition of failure in this context. Failure of effective AI integration will consist in one or both of the following: a) when AI does not achieve performance comparable to or better than radiologists, when evaluated against the same reference standard, b) when AI introduces added costs or inefficiencies, including delays or workflow disruptions. The group members independently hand-searched recent literature looking for verbal or conceptual synonyms of failure within AI implementation (such as error, mistake, inefficiency, ineffectiveness, shortcoming, problem, challenge, delay, barrier, "reduced accuracy", "does not work", etc) in titles or abstracts and brought together published peer-reviewed papers which documented these AI failures in Radiology. Similarly, papers including solutions to these errors were searched. All eligible papers were then analysed by 3 group members, who assigned them to one of the 3 components of AI implementation using consensus, and then compiled a complete draft based on these papers, which was live edited by the wider team. This paper offers the "bigger picture" of AI implementation errors and solutions in Radiology, addressing all essential components, as described above, and presented as a comprehensive narrative review of published research evidence. This is not an exhaustive list of either AI implementation failures or their solutions. This is not merely because of the known publication bias. It is also because it is not only difficult to define failure itself in an ever-expanding field, but also challenging to search for failure in a culture that mostly celebrates success. Documenting and understanding these failures is the first step towards resolving them in the present and preventing them in the future, for optimising the use and impact of AI in radiology.

2. Reasons for AI implementation failures

2.1. Issues with AI models throughout the lifecycle

The AI model lifecycle consists of six independent stages, which are: i) inception of the AI solution, ii) development, iii) internal and external testing, iv) deployment, v) operation, maintenance (post-market surveillance and versioning) and re-evaluation and vi) decommissioning [13] (Fig. 1). Failures can be observed during all the above stages. These are described in detail further below.

2.1.1. Inception

Early stakeholder engagement—including patients and clinicians—is crucial during inception. Delays in engaging them can impact acceptability, usability and ergonomics, undermining effective implementation [14,15]. Their involvement is essential to ensure the model is genuinely driven by clinical needs (clinical pull), rather than shaped by technical convenience (technical push), as clinical scenarios are inherently complex and multifactorial [16]. Moreover, involving all stakeholders as early as possible enhances trust in AI models [14,17]. Trust can be further enhanced by identifying risks early from various perspectives (clinical, technical, legal, ethical, social, economic, etc) and proposing mitigations.

Ambiguity in defining intended use during the inception phase can create a misalignment with clinical needs and user expectations, potentially compromising patient safety and outcomes. Clear definition of intended use is therefore critical to avoid inappropriate or off-label use in unintended settings [18,19]. The intended use statement, a statement that developers are now required to provide in order to obtain Conformité Européenne (CE) mark, to comply with the EU AI Act or to obtain clearance from the US Food and Drug Administration (FDA), is still, in many cases, not publicly declared by manufacturers [20]. This can lead to misconceptions regarding the actual use of AI tools, and misuse of these solutions in clinical practice. In addition, this may lead to challenges in procurement of AI tools, increasing costs and deterring buyers.

2.1.2. AI development

AI tools can fail due to lack of robustness, limited generalisability and algorithmic bias. There are various types of algorithmic bias, often stemming from suboptimal training datasets [21] that inadequately represent patient subgroups across age, sex, race, and socioeconomic status [22–27]. Biases can arise from poor dataset composition, lack of fairness in model testing, and low data quality, potentially leading to selection bias, over- or underfitting for specific pathologies [28], and false predictions [25]. Consequently, model performance may drop when applied to previously unseen clinical data [29].

Ideally, datasets should be diverse and inclusive. Limitations of historical datasets, governance issues related to protected characteristics, and data sharing restrictions can hinder model performance. Gender bias, for instance, can impair recognition of sex-specific clinical presentations, leading to underdiagnosis and worse outcomes, including increased morbidity and mortality for affected groups [30]. Such biases should not be amplified by AI integration [16], as flawed scaling could exacerbate harm. A recent US review found that imaging AI data (2015–2019) came mostly from just three states [31], and the UK Biobank data contains only 6 % non-European ancestry participants [26]. Legal frameworks in some countries may also restrict inclusion of features like ethnicity or race [32]. Balancing datasets prior to deployment helps reduce bias and improve model performance and patient outcomes [28]. Federated learning—training models across institutions without direct data sharing—offers a solution for building diverse models under stricter data governance regimes, assuming population variability across institutions [26].

While large datasets are traditionally favoured in AI development for generalisability, size alone isn't sufficient, because quality of data is equally important. Some studies relying on large datasets may report lower diagnostic accuracy than those with smaller ones [33].

2.1.3. AI testing

Testing on both internal and external data has been widely used when developing AI algorithms in radiology to ascertain the robustness of their performance [21,34,35]. An AI model may have achieved high performance when tested on internal data, which has substantially decreased when faced with real-world unseen data [36]. Recent research has confirmed a decrease in AI tools' algorithmic performance in radiology of up to 24 %, when external testing was applied [37]. External testing can ensure robust model performance, reproducibility and generalisability when performing beyond the training cohort [38]. When performing external testing prior to clinical integration, it is crucial to: judiciously define the ground truth (or reference standard), avoid data contamination or leakage (training data kept separate to testing data), test large enough datasets to ensure statistical power for a given context, test appropriate, diverse use cases which reflect the actual patient population on which the solution will run, and choose the appropriate evaluation metrics [33,39].

A major challenge with AI testing is that it is often of limited scope. Typically testing assesses accuracy and sometimes generalisability using external testing, but not always explores fairness, robustness, safety, usability, productivity, acceptance, explainability. That is why failures can occur even when appropriate external testing has been performed, and the model may prove unreliable even when ran across multiple clinical sites without careful consideration of the population demographics [40]. Therefore, testing on benchmarking external datasets across multiple sites is preferable [41], including sub-analysis of the results, to ensure consistently high performance and interoperability across different clinical contexts [42]. Employing specific frameworks to evaluate AI tools, including specific questions around intended use, is also important for testing [43]. Multifaceted testing beyond accuracy can also prevent mistakes in implementation from the outset.

2.1.4. Deployment of AI

Medical imaging AI solutions are being increasingly deployed and used in daily practice globally [1,44,45]. At the real-life deployment stage factors may arise that have not been encountered at the development or testing phase. This could relate to variations in real-world data, such as anatomical variations [46–49], distinct, new or rare pathologies [50–52], atypical uptake of contrast media [53], complex post-operative findings [28], foreign objects [54], and image quality variations (image rotation, motion and other artefacts, differences in image brightness, resolution, and contrast) [55].

The performance of AI models may also be affected by the increasing complexity of clinical cases. For instance, the accuracy of AI tools applied for fracture detection in spine CT scans can be negatively affected by the type of pathology (e.g. presence of chronic fractures, osteophyte formation), their visibility on sagittal series and patient's age [24]. In many cases algorithms performed worse than radiologists, despite detecting most fractures that were undetected by them [19] due to the complexity of the specific clinical context (in this case, vertebral fractures). Similar results have been observed when applied for the detection of major pathologies in head CT scans, where the type and number of brain haemorrhages and prior neurosurgery [56], or where location of large vessel occlusion significantly affected the model's accuracy [57].

Similarly, when using real-world cohorts, the performance of certain AI methods, such as radiomics feature extraction, can be considerably lower, as recent research on prediction and survival response indicates [58]. This has been also confirmed by recent research attempting to produce normative brain volume reports from real-world data [59]. Varying model sensitivities have been also observed depending on the detection probability threshold used by researchers [60].

2.1.5. Post-market surveillance

Failures may occur long after initial deployment, due to innate changes in the performance of AI algorithms (called model drift) [61], equipment (software or hardware) changes or upgrades, the different personnel using the AI tools, the evolving patient demographics (data drift) [62], and the emergence of new pathologies (e.g. new virus strains or epidemics) that require continuously new diagnostic approaches [63]. The importance of documenting these failures can be reflected on the FDA Total Product Life Cycle approach, which requires real-world performance monitoring for any software used as a medical device (SaMD) [64]. In addition, post-market surveillance and human oversight are a requirement both by the recent EU AI Act [65] and the Medical Device Regulation [65,66]. Both periodic evaluation of the model's performance by institutions and post-market surveillance performed by industry have been recommended but are often challenging to implement and are not always performed by AI vendors [67]. AI vendors support their value propositions, by promising their product could lead to cost reduction, workflow efficiency, workload reduction, and many more claims. However, it is now evident that many largely fail to provide sustained evidence on how their products meet clinical needs in the longer-term and deliver these values in clinical practice [68].

Continuous monitoring after deployment is also crucial in cases where adaptive clinical decision systems (CDS) have been implemented into clinical practice, since these decision support tools dynamically train themselves instead of relying on predefined data training, and this might naturally cause variations in the performance [69].

2.1.6. Product decommissioning

Another important aspect of the AI product lifecycle is decommissioning. This can occur at any point due to continued model inaccuracies, induced workflow complexities, high or non-manageable maintenance costs [70,71]. During decommissioning, potential challenges relate to safely managing any stored/archived data within the AI tool, ceasing accessibility of the vendor to further clinical data through pipelines set-up previously, and aligning with any local or national data storage/security requirements [72]. Importantly, AI governance principles must be followed throughout the lifecycle, from inception to product decommissioning [73].

2.2. Infrastructure issues

Another source of potential failures is related to AI infrastructure. Infrastructure relates to all hardware and software needed to develop and deploy AI solutions [74]. This includes the baseline Information Technology (IT) infrastructure but also extends to infrastructure relating to software and hardware designed to serve AI pipelines only. Most healthcare organisations build their own AI infrastructure which, as a bare minimum, includes on-premises computing and storage. However, because of often complex governance and the associated large up-front financial investment required, these systems, where available, remain largely segregated from those used in other clinical settings, reducing the possibility of continuity of care or healthcare data integration that could benefit both patients and organisations [75]. Other clinical settings rely on inbuilt solutions, which, while easier to customise, cannot always be upgraded or maintained, as required for sustained impact.

2.2.1. Networks

A potential source of failures is related to the IT infrastructure of the clinical site and the strategies employed for AI implementation [75]. AI failures could be triggered either when integrating solutions locally on site, or when integrating them as a cloud-based software model. Proper data orchestration is crucial, and this should include Digital Imaging and Communications in Medicine (DICOM) and Health Level Seven (HL7) data [76] to ensure that the right data is sent to the appropriate AI tools. Moreover, interoperability across different infrastructure networks, to enhance and streamline information sharing, and a standards-based approach are to improve clinical integration [77].

2.2.2. Integration into picture archiving and communication system (PACS) and radiology information system (RIS)

The integration of AI solutions into the PACS of clinical sites represents an additional challenge. Most PACS currently allow only one way data transfer. This may result in misinterpretations by the referring clinicians, who have access to PACS, but may have not yet been trained to evaluate AI-enabled results from medical imaging [78]. Radiologists need to be given the option to modify the results before publication, interact with AI, and store the modified results into PACS. This could allow an interaction that would empower clinicians to reprocess data and generate new outcomes. Also, user-friendliness of AI infrastructure could involve a push-to-PACS function, so that radiologists do not have to manually retrieve AI results each time [79]. Seamless integration with both PACS and RIS is vital to minimise further errors between imaging data and patient scheduling [80]. With upcoming AI models producing a (partial) draft of the radiology report (report pre-population), it is important to integrate these pre-populated reports into the reporting system (RIS or other) and to allow radiologists to confirm or modify these accordingly. Many of these challenges can be overcome by the use of unified platforms that may enable software of different vendors into a central docking system for better integration or employ the widget approach, that allows the user to select AI solutions from a unified interface [81].

2.2.3. Hardware

The use of high-performance hardware is a requirement for AI functionality, for safe and efficient processing, storage and management of vast amounts of data. While central processing units (CPUs) serve as general-purpose processors and the core of computing systems, AI tasks typically require the parallel processing power of graphics processing units (GPUs), originally designed for rendering graphics but now essential for AI applications [82]. GPUs handle large data volumes efficiently and outperform CPUs in most AI tasks, including MRI image analysis [83,84]. Hardware used for AI in radiology should, therefore, have the capacity (memory and processing power) to perform complex computational processes simultaneously, while software used should allow seamless integration into existing workflows [85]. However, GPUs are costly and are in limited supply, with one company controlling around 80 % of the global GPU market [86,87]. This scarcity can compromise AI performance due to suboptimal computational resources. It must be noted that there is limited published research on the failures observed due to infrastructure challenges/deficiencies in radiology. With more clinical sites moving forward with AI deployment and AI models becoming more computationally intensive, this topic deserves more attention and ongoing research.

2.3. Human factors

As discussed in introduction, this is the third component of AI implementation success and/or failure. We will look at its different constituents, physical, cognitive, perceptual, emotional, procedural, and sociocultural, and how these might challenge AI use and deployment.

2.3.1. Socio – Cultural aspects

Social and cultural aspects of AI implementation and societal values and preferences should be considered when designing, training, deploying, and evaluating AI tools, to ensure they address local needs, mitigate perceived risks and harness benefits for the respective sociocultural context [88].

2.3.1.1. Resistance to change. It is natural for healthcare professionals to resist change, as it demands adjustments to established routines in high-stakes environments where time and resources are limited. This resistance to culture change may be a major impediment for AI implementation [89]. Successful deployment requires substantial changes in the work habits, mindset, and organisational aspects of AI adopters [90]. End-users should consider the ethical challenges derived from AI integration [91], such as data privacy, model performance, patient safety and societal impact (e.g. sustainability, cost etc), and employ formal change management strategies, such as inclusion in the decision making and coproduction, to minimise resistance to change and foster long-lasting, meaningful AI innovations [90]. The importance of adopting change management strategies, and considering the potential failures that these may incur, is further exacerbated by research showing an inherent resistance to change among healthcare systems and healthcare professionals [92,93]. In essence, drastic change of current practices is necessary and more important than the attempt to incorporate AI into decade-old practices.

2.3.1.2. Publication mishaps, reporting discrepancies, and a black-box culture. As mentioned in introduction, publication bias represents a significant challenge for educating practitioners, patients and the public around AI. Withholding information about failing AI systems, processes and practices can mislead clinicians and the public to believe that AI is infallible, or that we need not question its results, leading to AI overreliance [4]. This lack of adequate published evidence on AI failures may misrepresent and underplay the real-world challenges associated with AI adoption, leaving radiologists and other healthcare professionals exposed to unnecessary risks and duplication of effort. It also creates unrealistic expectations of error-free AI, undermines the value of human intelligence as a critical actor to AI evaluation, erodes the foundations of human-AI collaborative work [94], may also create a culture of blame, denial and algorithmic aversion in more experienced readers, and result into suboptimal integration in clinical practice, with direct impact on patient outcomes [95,96]. Another important drawback associated with AI-related publications refers to the black-box culture and the associated lack of open-source code provision in manuscripts, which prevents testing of the reproducibility of their results [97,98]. In addition, evaluation of different AI models in the literature has shown that use of metrics is often haphazard, lacks scientific justification, and, in many cases fails to provide the readers with the true performance of the model [33]. It should be noted that better algorithmic performance metrics does not necessarily translate to optimal clinical outcomes. Many radiology studies have been designed as retrospective cohort studies, having limited external validation and being prone to bias. In addition, most of them use a narrow range of techniques and selection of cases [99]. Finally, despite the considerable increase in the AI-enabled Radiology clinical trials, many studies report findings from single-centre datasets, providing little information on demographic data, and report varying operational efficiency, thus their findings should be interpreted with caution [100,101].

Any AI failures must be adequately reported by researchers, and appropriate steps must be taken, using multi-institutional, multinational, multimodal datasets to this direction [5]. Supporting these observations, a recent umbrella review assessing adherence to the Checklist for Artificial Intelligence in Medical Imaging (CLAIM)—currently the highest level of evidence on the topic—identified the item related to failure analysis as one of the 11 most underreported items [102]. Notably, more than 75 % of AI studies failed to report this critical information. As mandated by the CLAIM guidelines, authors should provide detailed information about algorithmic failures. Unfortunately, true adherence to CLAIM is still low, as authors fail to report all required details in their manuscripts, despite reference to the CLAIM checklist in their methods [102].

2.3.2. Physical aspects and ergonomics

A vital aspect of AI implementation is our ability to smoothly integrate these tools into clinical practice while preserving efficiency. This includes optimal strategies to ensure a good level of ergonomics, since physical environment is crucial for professionals' efficiency and efficacy [103]. Errors could increase if users are required to become familiar with different, complex interfaces, and this could also limit user acceptance [78]. The interface employed by AI tools should be simple, user-friendly, accessible, and ideally fully integrated into one platform [104,105] within the reporting environment, otherwise image reporting and other tasks may be too laborious for adoption.

2.3.3. Emotional and perceptual challenges, human-AI interaction and automation bias

If not thoughtfully and holistically integrated, the daily use of AI tools in clinical practice could impact emotional wellbeing of end-users. Recent studies report that radiologists using AI solutions exhibited higher levels of burnout compared to those who did not use AI in practice due to workflow complexities, lack of training, and impact on decision making [106]. This, in turn, might cause radiologists to be prone to mistakes during their daily practice.

Optimising human-AI interaction is crucial for improving system performance and ensuring clinician well-being. While autonomous AI can sometimes outperform human-AI collaboration in specific use cases [107], its expanding use within multidisciplinary teams may introduce tensions and power imbalances during handovers [108]. Building trust and transparency for the interaction between practitioners and AI is essential for seamless collaboration. Enhancing human situation awareness is also key to anticipating and mitigating the impact of changes in real time [12].

Suboptimal human-AI interaction can lead to reduced efficiency, such as increased reading times [106,109], and performance issues due to systematic errors from overreliance on AI [95]. Studies show that incorrect AI outputs can raise false positive and negative rates among radiologists [31,109,110–112], with less experienced clinicians being more susceptible to automation bias [6,113]. In some cases, AI may impair radiologists' ability to detect both normal and pathological findings, potentially affecting treatment and patient outcomes [114].

2.3.4. Cognitive aspects and annotation or interpretation errors

Data labelling by human experts is required for supervised learning training of AI models in medical imaging. Annotators use labelling to mark medical images as normal or abnormal, with different pathologies, so that the AI algorithms can then be trained to discriminate them. However, labelling mistakes resulting from differing levels of observer experience, cognitive fatigue, lack of attention, or reduced concentration can cause AI inaccuracies [33]. The same is true for image interpretation tasks. Incomplete or inconsistent labelling causes reproducibility concerns for the AI model and those who use it [5]. Employing self-supervised learning approaches for image annotation tasks could serve as a remedy for lack of reproducibility [115], with recent research showing that these strategies can outperform all supervised methods [116]. Extending this paradigm further to image interpretation, AI combining computer vision and natural language processing has been proposed as the ultimate quality assurance tool to review a radiologist's report, identifying discrepancies between medical imaging and documentation [117].

2.3.5. Procedural aspects

2.3.5.1. Workflow efficiency: An AI promise still waiting to be fulfilled. One of the biggest promises of AI in medical imaging is its potential to improve efficiency of clinical workflows and reduce turnaround times, critical at a time when waiting lists are expanding and staffing shortages are increasing [118]. In this context, it is expected that AI tools will help radiologists reduce their cognitive load and dedicate less time per examination. However, this is not always the case; certain AI tools may lengthen reading times of radiologists [119], for instance when normal examinations are falsely flagged as pathologic [18]. The implementation of AI tools in clinical practice may not therefore always realise the potential for increased patient throughput. Their minimal impact on streamlining clinical workflows [120], without observed differences in the workload and stress of radiologists may suggest that adoption of AI tools may occasionally create new challenges for medical imaging professionals [27,110]. Important to ensure any performance evaluation window is long enough to allow ample time for human and AI interaction before drawing any definitive conclusions. This underlines the need of post market surveillance.

2.3.5.2. Regulation and policy. There is currently a wealth of regulatory frameworks globally. While regulation plays a crucial role in the implementation of AI, as it can ensure safe use and monitoring of AI tools in clinical practice, its complexity may also prove challenging for deployment in clinical contexts. Robust regulation can prevent failures caused due to suboptimal data protection, insufficient testing of AI algorithms, and help manage accountability concerns [121]. Regulatory frameworks must be optimised in national contexts and harmonised at an international level, covering the whole AI lifecycle in healthcare. Failures due to an algorithmic drift should be monitored with robust post market surveillance, and algorithms updated accordingly under pre-determined change control plans [122].

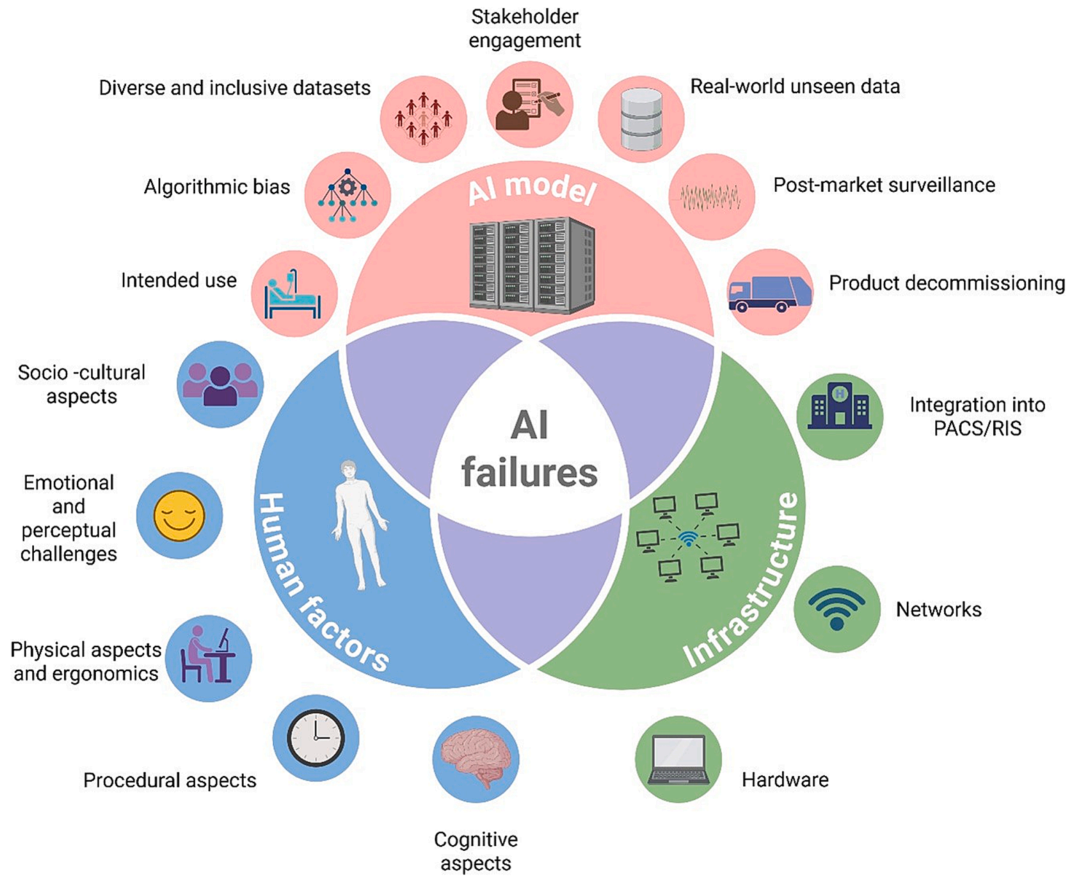

Fig. 2. AI failures related to human factors, models and infrastructure (Created with Biorender).

3. What can we do about it?

In the above exploration of AI failures (see Fig. 2 for a holistic view) specific remedial action was proposed when supported by respective research evidence. This section proposes some more universal solutions to AI errors, from both published evidence and own expertise.

3.1. AI education and literacy

Appropriate AI education is a key step to prevent, recognise or correct AI failures within radiology. Training should focus not only on the theoretical principles of AI and its applications, but also explicitly discuss the sources/mechanisms of failures [4]. This educational provision should support the development of essential perceptual skills, critical reflection, knowledge and understanding of related governance and appropriate use of AI without overreliance on technology. Customised training can enable radiologists and other healthcare professionals anticipate, recognise, and manage AI failures in practice. As outlined in the EU AI Act, qualified professionals must provide human oversight after AI deployment [123]. The vital role of AI literacy for successful implementation has been widely discussed within the medical imaging field [124,125], and specific digital competencies are now a requirement for radiographers to practise in the UK [126]. Moreover, AI literacy has been widely promoted by key professional bodies and statements [65]. Formal, personalised, academically-accredited educational provisions employing an interprofessional approach [127,128] will help practitioners gain deeper understanding of AI failures and safely navigate the new AI era.

Medical imaging professionals should view AI failures as learning opportunities to improve algorithms and deployment strategies. Creating an open repository of AI erroneous cases with analyses of causes and suggestions for improvement, would be invaluable for training future professionals, while enhancing human-AI interaction. Moreover, future AI development could be accelerated through repositories enabling crowdsourced testing by individuals with diverse expertise [129].

3.2. Continuous monitoring of AI

As previously mentioned, many AI failures occur after the deployment in clinical practice, and these can be often overlooked by end-users if evaluation occurs only at one timepoint. Post-market surveillance, a continuous evaluation of AI systems performance, is reinforced by the EU AI Act, mandating that AI vendors develop specific monitoring systems to continuously assess AI performance after clinical integration [130]. Continuous monitoring of deployed AI systems will also enhance practitioner and patient trust in AI technologies and help maintain high standards of care delivery [65]. Research dissemination and published evidence of AI tools' performance should also report longitudinal evaluations of these tools in clinical practice.

3.3. Standardising reporting of AI studies

CLAIM offers a straightforward, standardised way for researchers to ensure consistent reporting of AI-related studies in medical imaging [34]. Similarly, researchers can also benefit from other important guidelines for standardised reporting of prediction model studies, such as the recently updated Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis Artificial Intelligence (TRIPOD + AI) statement [131] or Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence (DECIDE AI) [132] for the early-stage clinical evaluation of AI-driven decision support systems. Finally, the SPIRIT AI AND CONSORT AI guidelines [133] should be used instead for standardizing the reporting of randomized clinical trials.

3.4. Multiprofessional collaboration and AI leadership

To effectively manage integration of AI into clinical practice, and minimise failures, it is important to build strong, collaborative AI teams. Collaboration should focus on the inclusion of all key stakeholders of the medical imaging AI ecosystem, adopting a multidisciplinary approach to harness the benefits of AI and mitigate any potential risks [134]. Thus, it will be possible to gain more insights into areas where understanding or research is lacking, like AI infrastructure, as described above. Furthermore, effective leadership is required to bring together and coordinate all different aspects of AI implementation (people, processes, product, and technology) [135] to maximise success of AI solutions despite the complexities of clinical practice. This is strengthened by recent research highlighting leadership errors as the most important cause of AI failures within companies [136].

3.5. Funding and business cases

Finally, financial constraints within organisations may result in premature AI project termination and incomplete implementation. Optimal allocation of financial resources, robust business cases, and proper reimbursement of healthcare systems will ensure the necessary capital investment for infrastructure, training of professionals and recruitment of experts required to realise the promised organisational efficiencies and clinical efficacy, and ensure the longevity of AI deployment and monitoring [137]. More recently the FUTURE.AI project consortium proposed specific solutions for different AI challenges, which could offer a roadmap for streamlining AI implementation across healthcare and medical imaging, in particular [42].

4. Conclusion

AI can revolutionise medical imaging and other healthcare disciplines. The success of its implementation, though, relies on many factors; the knowledge and understanding of its potential errors and inadequacies; the lifelong learning of human observers so they can reskill to address these challenges; the collaborative spirit of multidisciplinary teams guided by open-minded leaders; the careful coordination of processes, people, products and infrastructure; and a culture change, where failure is celebrated as an opportunity to learn and expand our understanding of the capabilities of AI for human benefit.

CRediT authorship contribution statement

Nikolaos Stogiannos: Investigation, Visualization, Writing – original draft. **Renato Cuocolo:** Data curation, Writing – review & editing. **Tugba Akinci D'Antonoli:** Data curation, Writing – review & editing. **Daniel Pinto dos Santos:** Data curation, Writing – review & editing. **Hugh Harvey:** Data curation, Writing – review & editing. **Merel Huisman:** Data curation, Writing – review & editing. **Burak Kocak:** Data curation, Writing – review & editing. **Elmar Kotter:** Data curation, Writing – review & editing. **Karim Lekadir:** Data curation, Writing – review & editing. **Susan Cheng Shelmerdine:** Data curation, Writing – review & editing. **Kicky G van Leeuwen:** Data curation. **Peter van Ooijen:** Data curation. **Michail E. Klontzas:** Conceptualization, Data curation, Supervision, Writing – review & editing. **Christina Malamateniou:** Conceptualization, Data curation, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T.G. Day, J. Matthew, S.F. Budd, L. Venturini, R. Wright, A. Farruggia, T. V. Vigneswaran, V. Zidere, J.V. Hajnal, R. Razavi, J.M. Simpson, B. Kainz, Interaction between clinicians and artificial intelligence to detect fetal atrioventricular septal defects on ultrasound: how can we optimize collaborative performance? Ultrasound Obstet. Gynecol. 64 (2024) 28–35, https://doi.org/10.1002/uog.27577.
[2] M. Vaccaro, A. Almaatouq, T. Malone, When combinations of humans and AI are useful: a systematic review and meta-analysis, Nat. Hum. Behav. 8 (2024) 2293–2303, https://doi.org/10.1038/s41562-024-02024-1.
[3] E.J. Topol, High-performance medicine: the convergence of human and artificial intelligence, Nat. Med. 25 (2019) 44–56, https://doi.org/10.1038/s41591-018-0300-7.
[4] M.D. Li, B.P. Little, Appropriate reliance on artificial intelligence in radiology education, J. Am. Coll. Radiol. 20 (2023) 1126–1130, https://doi.org/10.1016/j.jacr.2023.04.019.
[5] S. Purkayastha, H. Trivedi, J.W. Gichoya, Failures hiding in success for artificial intelligence in radiology, J. Am. Coll. Radiol. 18 (2021) 517–519, https://doi.org/10.1016/j.jacr.2020.11.008.
[6] T. Dratsch, X. Chen, M. Rezazade Mehrizi, R. Kloeckner, A. Mähringer-Kunz, M. Püsken, B. Baeßler, S. Sauer, D. Maintz, D. Pinto Dos Santos, Automation bias in mammography: the impact of artificial intelligence BI-RADS suggestions on reader performance, Radiology 307 (2023) e222176, https://doi.org/10.1148/radiol.222176.
[7] S. Mo Jones-Yang, Y.J. Park, How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability, JCMC 28 (2023) 1–8, https://doi.org/10.1093/jcmc/zmac029.
[8] B. Kocak, E. Bulut, O.N. Bayrak, A.A. Okumus, O. Altun, Z. Borekci Arvas, I. Kavukoglu, NEgatiVE results in Radiomics research (NEVER): a meta-research study of publication bias in leading radiology journals, Eur. J. Radiol. 163 (2023) 110830, https://doi.org/10.1016/j.ejrad.2023.110830.
[9] B. Kocak, D. Pinto dos Santos, M. Dietzel, The widening gap between radiomics research and clinical translation: rethinking current practices and shared responsibilities, EJR AI 1 (2025) 100004, https://doi.org/10.1016/j.ejrai.2025.100004.
[10] A.S. Nair, Publication bias – importance of studies with negative results!, Indian J. Anaesth. 63 (2019) 505–507, https://doi.org/10.4103/ija.IJA_142_19.
[11] D. Kreuzberger, N. Kühl, S. Hirschl, Machine learning operations (MLOps): overview, definition, and architecture, IEEE Access 11 (2023) 31866–31879, https://doi.org/10.1109/ACCESS.2023.3262138.
[12] M. Sujan, D. Furniss, K. Grundy, H. Grundy, D. Nelson, M. Elliott, S. White, I. Habli, N. Reynolds, Human factor challenges for the safe use of artificial intelligence in patient care, BMJ Health Care Inform. 26 (2019) e100081, https://doi.org/10.1136/bmjhci-2019-100081.
[13] M. Sujan, C. Smith-Frazer, C. Malamateniou, J. Connor, A. Gardner, H. Unsworth, H. Husain, Validation framework for the use of AI in healthcare: overview of the new British standard BS30440, BMJ Health Care Inform. 30 (2023) e100749, https://doi.org/10.1136/bmjhci-2023-100749.
[14] S. Banerjee, P. Alsop, L. Jones, R.N. Cardinal, Patient and public involvement to build trust in artificial intelligence: a framework, tools, and case studies, Patterns (n y) 3 (2022) 100506, https://doi.org/10.1016/j.patter.2022.100506.
[15] S. Adus, J. Macklin, A. Pinto, Exploring patient perspectives on how they can and should be engaged in the development of artificial intelligence (AI) applications in health care, BMC Health Serv. Res. 23 (2023) 1163, https://doi.org/10.1186/s12913-023-10098-2.
[16] E. Dahlin, Mind the gap! on the future of AI research, Humanit. Soc. Sci. Commun. 8 (2021) 71, https://doi.org/10.1057/s41599-021-00750-9.
[17] H. Xu, K.M.J. Shuttleworth, Medical artificial intelligence and the black box problem: a view based on the ethical principle of "do no harm", Intell. Med. 4 (2024) 52–57, https://doi.org/10.1016/j.imed.2023.08.001.
[18] A.J. Del Gaizo, T.F. Osborne, T. Shahoumian, R. Sherrier, Deep learning to detect intracranial hemorrhage in a national teleradiology program and the impact on interpretation time, Radiol. Artif. Intell. 6 (2024) e240067, https://doi.org/10.1148/ryai.240067.
[19] G.J. van den Wittenboer, B.Y.M. van der Kolk, I.M. Nijholt, E. Langius-Wiffen, R. A. van Dijk, B.A.A.M. van Hasselt, M. Podlogar, W.A. van den Brink, G.J. Bouma, N.W.L. Schep, M. Maas, M.F. Boomsma, Diagnostic accuracy of an artificial intelligence algorithm versus radiologists for fracture detection on cervical spine CT, Eur. Radiol. 34 (2024) 5041–5048, https://doi.org/10.1007/s00330-023-10559-6.
[20] K.G. van Leeuwen, D.M. Hedderich, H. Harvey, S. Schalekamp, How AI should be used in radiology: assessing ambiguity and completeness of intended use statements of commercial AI products, Insights Imaging 15 (2024) 51, https://doi.org/10.1186/s13244-024-01616-9.
[21] A. Maiter, K. Hocking, S. Matthews, J. Taylor, M. Sharkey, P. Metherall, S. Alabed, K. Dwivedi, Y. Shahin, E. Anderson, S. Holt, C. Rowbotham, M. A. Kamil, N. Hoggard, S.P. Balasubramanian, A. Swift, C.S. Johns, Evaluating the performance of artificial intelligence software for lung nodule detection on chest radiographs in a retrospective real-world UK population, BMJ Open 13 (2023) e077348, https://doi.org/10.1136/bmjopen-2023-077348.

[22] H.S. Oh, T.H. Kim, J.W. Kim, J. Yang, H.S. Lee, J.H. Lee, C.H. Park, Feasibility and limitations of deep learning-based coronary calcium scoring in PET-CT: a comparison with coronary calcium score CT, Eur. Radiol. 34 (2024) 4077–4088, https://doi.org/10.1007/s00330-023-10390-z.

[23] L. Seyyed-Kalantari, H. Zhang, M.B.A. McDermott, I.Y. Chen, M. Ghassemi, Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations, Nat. Med. 27 (2021) 2176–2182, https://doi.org/10.1038/s41591-021-01595-0.

[24] A.F. Voter, M.E. Larson, J.W. Garrett, J.J. Yu, Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures, AJNR Am. J. Neuroradiol. 42 (2021) 1550–1556, https://doi.org/10.3174/ajnr.a7179.

[25] B. Glocker, C. Jones, M. Roschewitz, S. Winzeck, Risk of bias in chest radiography deep learning foundation models, Radiol. Artif. Intell. 5 (2023) e230060, https://doi.org/10.1148/ryai.230060.

[26] J.W. Gichoya, K. Thomas, L.A. Celi, N. Safdar, I. Banerjee, J.D. Banja, L. Seyyed-Kalantari, H. Trivedi, S. Purkayastha, AI pitfalls and what not to do: mitigating bias in AI, Br. J. Radiol. 96 (2023) 20230023, https://doi.org/10.1259/bjr.20230023.

[27] F.P. Schweikhard, A. Kosanke, S. Lange, M.L. Kromrey, F. Mankertz, J. Gamain, M. Kirsch, B. Rosenberg, N. Hosten, Doctor's orders-why radiologists should consider adjusting commercial machine learning applications in chest radiography to fit their specific needs, Healthcare (Basel) 12 (2024) 706, https://doi.org/10.3390/healthcare12070706.

[28] A.P. Brady, E. Neri E, 2020. Artificial Intelligence in Radiology-Ethical Considerations. Diagnostics (Basel) 10, 231. doi: 10.3390/diagnostics10040231.

[29] S. Katal, B. York, A. Gholamrezanezhad, AI in radiology: from promise to practice - a guide to effective integration, Eur. J. Radiol. 181 (2024) 111798, https://doi.org/10.1016/j.ejrad.2024.111798.

[30] T.Y. Sun, J. Hardin, H.R. Nieva, K. Natarajan, R.F. Cheng, P. Ryan, N. Elhadad, 2023. Large-scale characterization of gender differences in diagnosis prevalence and time to diagnosis. medRxiv [Preprint]. doi: 10.1101/2023.10.12.23296976.

[31] A. Kausal, R. Altman, C. Langlotz, Geographic distribution of US Cohorts used to train deep learning algorithms, J. Am. Med. Assoc. 324 (2020) 1212–1213, https://doi.org/10.1001/jama.2020.12067.

[32] European Union Agency for Fundamental Rights. Second European Union Minorities and Discrimination Survey - Main results. https://fra.europa.eu/en/publication/2017/second-european-union-minorities-and-discrimination-survey-main-results Published December 6 2017. Accessed March 17, 2025.

[33] G. Varoquaux, V. Cheplygina, Machine learning for medical imaging: methodological failures and recommendations for the future, NPJ Digital Med. 5 (2022) 48, https://doi.org/10.1038/s41746-022-00592-y.

[34] A.S. Tejani, M.E. Klontzas, A.A. Gatti, J.T. Mongan, L. Moy, S.H. Park, C. E. Kahn Jr, Checklist for artificial intelligence in medical imaging (CLAIM): 2024 update, Radiol. Artif. Intell. 6 (2024) e240300, https://doi.org/10.1148/ryai.240300.

[35] B. Kocak, T. Akinci D'Antonoli, N. Mercaldo, A. Alberich-Bayarri, B. Baessler, I. Ambrosini, A.E. Andreychenko, S. Bakas, R.G.H. Beets-Tan, K. Bressem, I. Buvat, R. Cannella, L.A. Cappellini, A.U. Cavallo, L.L. Chepelev, L.C.H. Chu, A. Demircioglu, N.M. deSouza, M. Dietzel, S.C. Fanni, A. Fedorov, L.S. Fournier, V. Giannini, R. Girometti, K.B.W. Groot Lipman, G. Kalarakis, B.S. Kelly, M. E. Klontzas, D.M. Koh, E. Kotter, H.Y. Lee, M. Maas, L. Marti-Bonmati, H. Müller, N. Obuchowski, F. Orlhac, N. Papanikolaou, E. Petrash, E. Pfaehler, D. Pinto Dos Santos, A. Ponsiglione, S. Sabater, F. Sardanelli, P. Seeböck, N.M. Sijtsema, A. Stanzione, A. Traverso, L. Ugga, M. Vallières, L.V. van Dijk, J.J.M. van Griethuysen, R.W. van Hamersvelt, P. van Ooijen, F. Vernuccio, A. Wang, S. Williams, J. Witowski, Z. Zhang, Z. Zwanenburg, R. Cuocolo, METhodological RadiomICs score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII, Insights Imaging 15 (2024) 8, https://doi.org/10.1186/s13244-023-01572-w.

[36] C. Mello-Thoms, C.A.B. Mello, Clinical applications of artificial intelligence in radiology, Br. J. Radiol. 96 (2023) 20221031, https://doi.org/10.1259/bjr.20221031.

[37] J.H.J. Ketola, S.I. Inkinen, T. Mäkelä, S. Syväranta, J. Peltonen, T. Kaasalainen, M. Kortesniemi, Testing process for artificial intelligence applications in radiology practice, Phys. Med. 128 (2024) 104842, https://doi.org/10.1016/j.ejmp.2024.104842.

[38] D. Daye, W.F. Wiggins, M.P. Lungren, T. Alkasab, N. Kottler, B. Allen, C.J. Roth, B.C. Bizzo, K. Durniak, J.A. Brink, D.B. Larson, K.J. Dreyer, C.P. Langlotz, Implementation of clinical artificial intelligence in radiology: who decides and how? Radiology 305 (2022) E62, https://doi.org/10.1148/radiol.229021.

[39] W. Tanguay, P. Acar, B. Fine, M. Abdolell, B. Gong, A. Cadrin-Chênevert, C. Chartrand-Lefebvre, J. Chalaoui, A. Gorgos, A.S. Chin, J. Prénovault, F. Guilbert, L. Létourneau-Guillon, J. Chong, A. Tang, Assessment of radiology artificial intelligence software: a validation and evaluation framework, Can. Assoc. Radiol. J. 74 (2023) 326–333, https://doi.org/10.1177/08465371221135760.

[40] A. Youssef, M. Pencina, A. Thakur, T. Zhu, D. Clifton, N.H. Shah, External validation of AI models in health should be replaced with recurring local validation, Nat. Med. 29 (2023) 2686–2687, https://doi.org/10.1038/s41591-023-02540-z.

[41] N. Sourlos, R. Vliegenthart, J. Santinha, M.E. Klontzas, R. Cuocolo, M. Huisman, P. van Ooijen, Recommendations for the creation of benchmark datasets for reproducible artificial intelligence in radiology, Insights Imaging 15 (2024) 248, https://doi.org/10.1186/s13244-024-01833-2.

[42] K. Lekadir, A.F. Frangi, A.R. Porras, B. Glocker, C. Cintas, C.P. Langlotz, E. Weicken, F.W. Asselbergs, F. Prior, G.S. Collins, G. Kaissis, G. Tsakou, I. Buvat, J. Kalpathy-Cramer, J. Mongan, J.A. Schnabel, K. Kushibar, K. Riklund, K. Marias, L.M. Amugongo, L.A. Fromont, L. Maier-Hein, L. Cerdá-Alberich, L. Martí-Bonmatí, M.J. Cardoso, M. Bobowicz, M. Shabani, M. Tsiknakis, M.A. Zuluaga, M. C. Fritzsche, M. Camacho, M.G. Linguraru, M. Wenzel, M. De Bruijne, M. G. Tolsgaard, M. Goisauf, M. Cano Abadía, N. Papanikolaou, N. Lazrak, O. Pujol, R. Osuala, S. Napel, S. Colantonio, S. Joshi, S. Klein, S. Aussó, W.A. Rogers, Z. Salahuddin, M.P.A. Starmans, FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare, BMJ 388 (2025) e081554, https://doi.org/10.1136/bmj-2024-081554.

[43] S.D. Khan, Z. Hoodbhoy, M.H.R. Raja, J.Y. Kim, H.D.J. Hogg, A.A.A. Manji, F. Gulamali, A. Hasan, A. Shaikh, S. Tajuddin, N.S. Khan, M.R. Patel, S. Balu, Z. Samad, M.P. Sendak, Frameworks for procurement, integration, monitoring, and evaluation of artificial intelligence tools in clinical settings: a systematic review, PLOS Digit. Health. 3 (2024) e0000514, https://doi.org/10.1371/journal.pdig.0000514.

[44] M. Zanardo, J.J. Visser, A. Colarieti, R. Cuocolo, M.E. Klontzas, D. Pinto Dos Santos, F. Sardanelli, Impact of AI on radiology: a EuroAIM/EuSoMII 2024 survey among members of the European Society of Radiology, Insights Imaging 15 (2024) 240, https://doi.org/10.1186/s13244-024-01801-w.

[45] S. Harris, T. Bonnici, T. Keen, W. Lilaonitkul, M.J. White, N. Swanepoel, Clinical deployment environments: five pillars of translational machine learning for health, Front. Digit. Health 4 (2022) 939292, https://doi.org/10.3389/fdgth.2022.939292.

[46] S. Behzad, S.M.H. Tabatabaei, M.Y. Lu, L.S. Eibschutz, A. Gholamrezanezhad, Pitfalls in Interpretive applications of Artificial intelligence in radiology, AJR Am. J. Roentgenol. 223 (2024) e2431493, https://doi.org/10.2214/ajr.24.31493.

[47] D.D. Martin, A.D. Calder, M.B. Ranke, G. Binder, H.H. Thodberg, Accuracy and self-validation of automated bone age determination, Sci. Rep. 12 (2022) 6388, https://doi.org/10.1038/s41598-022-10292-y.

[48] M. Ahluwalia, M. Abdalla, J. Sanayei, L. Seyyed-Kalantari, M. Hussain, A. Ali, B. Fine, The subgroup imperative: Chest radiograph classifier generalization gaps in patient, setting, and pathology subgroups, Radiol. Artif. Intell. 5 (2023) e220270, https://doi.org/10.1148/ryai.220270.

[49] L. Oakden-Rayner, W. Gale, T.A. Bonham, M.P. Lungren, G. Carneiro, A. P. Bradley, L.J. Palmer, Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study, Lancet Digit. Health 4 (2022) e351–e358, https://doi.org/10.1016/s2589-7500(22)00004-8.

[50] N. Alves, J.S. Bosma, K.V. Venkadesh, C. Jacobs, Z. Saghir, M. de Rooij, J. Hermans, H. Huisman, Prediction variability to identify reduced AI performance in cancer diagnosis at MRI and CT, Radiology 308 (2023) e230275, https://doi.org/10.1148/radiol.230275.

[51] N. Hasani, F. Farhadi, M.A. Morris, M. Nikpanah, A. Rhamim, Y. Xu, A. Pariser, M.T. Collins, R.M. Summers, E. Jones, E. Siegel, B. Saboury, Artificial intelligence in medical imaging and its impact on the rare disease community: threats, challenges and opportunities, PET Clin. 17 (2022) 13–29, https://doi.org/10.1016/j.cpet.2021.09.009.

[52] A. Aboshosha, AI based medical imagery diagnosis for COVID-19 disease examination and remedy, Sci. Rep. 15 (2025) 1607, https://doi.org/10.1038/s41598-024-84644-1.

[53] A. Hosny, D.S. Bitterman, C.V. Guthier, J.M. Qian, H. Roberts, S. Perni, A. Saraf, L.C. Peng, I. Pashtan, Z. Ye, B.H. Kann, D.E. Kozono, D. Christiani, P.J. Catalano, H.J.W.L. Aerts, R.H. Mak, Clinical validation of deep learning algorithms for radiotherapy targeting of non-small-cell lung cancer: an observational study, Lancet Digit. Health 4 (2022) e657–e666, https://doi.org/10.1016/s2589-7500(22)00129-7.

[54] D. Ng, H. Du, M.M. Yao, R.O. Kosik, W.P. Chan, M. Feng, Today's radiologists meet tomorrow's AI: the promises, pitfalls, and unbridled potential, Quant. Imaging Med. Surg. 11 (2021) 2775–2779, https://doi.org/10.21037/qims-20-1083.

[55] S.M. Santomartino, K. Putman, E. Beheshtian, V.S. Parekh, P.H. Yi, Evaluating the robustness of a deep learning bone age algorithm to clinical image variation using computational stress testing, Radiol. Artif. Intell. 6 (2024) e230240, https://doi.org/10.1148/ryai.230240.

[56] A.F. Voter, E. Meram, J.W. Garrett, J.J. Yu, Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of intracranial hemorrhage, J. Am. Coll. Radiol. 18 (2021) 1143–1152, https://doi.org/10.1016/j.jacr.2021.03.005.

[57] H. Mellander, A. Hillal, T. Ullberg, J. Wassélius, Evaluation of CINA® LVO artificial intelligence software for detection of large vessel occlusion in brain CT angiography, Eur. J. Radiol. Open 12 (2023) 100542, https://doi.org/10.1016/j.ejro.2023.100542.

[58] F. Peisen, A. Hänsch, A. Hering, A.S. Brendlin, S. Afat, K. Nikolaou, S. Gatidis, T. Eigentler, T. Amaral, J.H. Moltz, A.E. Othman, Combination of Whole-body baseline CT radiomics and clinical parameters to predict response and survival in a stage-IV melanoma cohort undergoing immunotherapy, Cancers (Basel) 14 (2022) 2992, https://doi.org/10.3390/cancers14122992.

[59] D.M. Hedderich, B. Schmitz-Koep, M. Schuberth, V. Schultz, S.J. Schlaeger, D. Schinz, C. Rubbert, J. Caspers, C. Zimmer, T. Grimmer, I. Yakushev, Impact of normative brain volume reports on the diagnosis of neurodegenerative dementia disorders in neuroradiology: a real-world, clinical practice study, Front. Aging Neurosci. 14 (2022) 971863, https://doi.org/10.3389/fnagi.2022.971863.

[60] X. Sun, T. Niwa, T. Okazaki, S. Kameda, S. Shibukawa, T. Horie, T. Kazama, A. Uchiyama, J. Hashimoto, Automatic detection of punctate white matter lesions

in infants using deep learning of composite images from two cases, Sci. Rep. 13 (2023) 4426, https://doi.org/10.1038/s41598-023-31403-3.

[61] R.E. Carter, V. Anand, D.M. Harmon Jr., P.A. Pellikka, Model drift: when it can be a sign of success and when it can be an occult problem, Intell-Based Med. 6 (2022) 100058, https://doi.org/10.1016/j.ibmed.2022.100058.

[62] B. Sahiner, W. Chen, R.K. Samala, N. Petrick, Data drift in medical machine learning: implications and potential remedies, Br. J. Radiol. 96 (2023) 20220878, https://doi.org/10.1259/bjr.20220878.

[63] J. Merkow, F.J. Dorfner, X. Yang, A. Ersoy, G. Dasegowda, M. Kalra, M.P. Lungren, C.P. Bridge, I. Tarapov. Scalable Drift Monitoring in Medical Imaging AI. Published 18 October 2024. [preprint]. https://arxiv.org/html/2410.13174v2.

[64] U.S. Food & Drug Administration. Total Product Life Cycle for Medical Devices. https://www.fda.gov/about-fda/cdrh-transparency/total-product-life-cycle-medical-devices Updated June 09 2023. Accessed April 11, 2025.

[65] E. Kotter, T.A. D'Antonoli, R. Cuocolo, M. Hierath, M. Huisman, M.E. Klontzas, L. Martí-Bonmatí, M.S. May, E. Neri, K. Nikolaou, D. Pinto Dos Santos, M. Radzina, S.C. Shelmerdine, A. Bellemo, Guiding AI in radiology: ESR's recommendations for effective implementation of the European AI Act, Insights Imaging 16 (2025) 33, https://doi.org/10.1186/s13244-025-01905-x.

[66] Medicines & Healthcare products Regulatory Agency. The Medical Devices (Post-market Surveillance Requirements) (Amendment) (Great Britain) Regulations 2024: guidance on implementation. https://www.gov.uk/government/publications/medical-devices-post-market-surveillance-requirements/the-medical-devices-post-market-surveillance-requirements-amendment-great-britain-regulations-2024-guidance-on-implementation Published 15 January 2025. Accessed April 11, 2025.

[67] The Royal College of Radiologists. Clinical Radiology: AI deployment fundamentals for medical imaging. https://www.rcr.ac.uk/media/sbdhwnfl/ai-deployment-fundamentals-for-medical-imaging-2024.pdf Published November 2024. Accessed March 07, 2025.

[68] M.H.R. Mehrizi, S.H. Gerritsen, W.M. de Klerk, C. Houtschild, S.M.H. Dinnessen, L. Zhao, R. van SomMehrimeren, A. Zerfu. How do providers of artificial intelligence (AI) solutions propose and legitimize the values of their solutions for supporting diagnostic radiology workflow? A technography study in 2021. Eur. Radiol. 33 (2023) 915-924. doi: 10.1007/s00330-022-09090-x.

[69] A.E. Solomonides, E. Koski, S.M. Atabaki, S. Weinberg, J.D. McGreevey, J. L. Kannry, C. Petersen, C.U. Lehmann, Defining AMIA's artificial intelligence principles, J. Am. Med. Inform. Assoc. 29 (2022) 585–591, https://doi.org/10.1093/jamia/ocac006.

[70] B.X. Collins, J.C. Bélisle-Pipon, B.J. Evans, K. Ferryman, X. Jiang, C. Nebeker, L. Novak, K. Roberts, M. Were, Z. Yin, V. Ravitsky, J. Coco, R. Hendricks-Sturrup, I. Williams, E.W. Clayton, B.A. Malin, Addressing ethical issues in healthcare artificial intelligence using a lifecycle-informed process, JAMIA Open 7 (2024) ooae108, https://doi.org/10.1093/jamiaopen/ooae108.

[71] K.G. van Leeuwen, M. de Rooij, S. Schalekamp, B. van Ginneken, M.J.C. M. Rutten, Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022, Eur. Radiol. 34 (2024) 348–354, https://doi.org/10.1007/s00330-023-09991-5.

[72] NHSx. A buyer's guide to AI in Health and Care. https://transform.england.nhs.uk/media/documents/NHSX_A_Buyers_Guide_to_AI_in_Health_and_Care.pdf Published November 2020. Accessed March 07, 2025.

[73] N. Stogiannos, R. Malik, A. Kumar, A. Barnes, M. Pogose, H. Harvey, M. F. McEntee, C. Malamateniou, Black box no more: a scoping review of AI governance frameworks to guide procurement and adoption of AI in medical imaging and radiotherapy in the UK, Br. J. Radiol. 96 (2023) 20221157, https://doi.org/10.1259/bjr.20221157.

[74] M. Flinders, I. Smalley. What is AI infrastructure? https://www.ibm.com/think/topics/ai-infrastructure#:~:text=AI%20infrastructure%20utilizes%20the%20latest,needed%20to%20train%20ML%20models Published June 3, 2024. Accessed March 07, 2025.

[75] T. Panch, H. Mattie, L.A. Celi, The "inconvenient truth" about AI in healthcare, NPJ Digit. Med. 2 (2019) 77, https://doi.org/10.1038/s41746-019-0155-4.

[76] A. Shah, P.S. Muddana, S. Halabi, A review of core concepts of imaging informatics, Cureus 14 (2022) e32828, https://doi.org/10.7759/cureus.32828.

[77] H. Kondylakis, V. Kalokyri, S. Sfakianakis, K. Marias, M. Tsiknakis, A. Jimenez-Pastor, E. Camacho-Ramos, I. Blanquer, J.D. Segrelles, S. López-Huguet, C. Barelle, M. Kogut-Czarkowska, G. Tsakou, N. Siopis, Z. Sakellariou, P. Bizopoulos, V. Drossou, A. Lalas, K. Votis, P. Mallol, L. Marti-Bonmati, L. C. Alberich, K. Seymour, S. Boucher, E. Ciarrocchi, L. Fromont, J. Rambla, A. Harms, A. Gutierrez, M.P.A. Starmans, F. Prior, J.L. Gelpi, K. Lekadir, Data infrastructures for AI in medical imaging: a report on the experiences of five EU projects, Eur. Radiol. Exp. 7 (2023) 20, https://doi.org/10.1186/s41747-023-00336-x.

[78] A.P. Brady, B. Allen, J. Chong, E. Kotter, N. Kottler, J. Mongan, L. Oakden-Rayner, D.P. Dos Santos, A. Tang, C. Wald, J. Slavotinek, 2024. Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. Insights Imaging 15, 16. doi: 10.1186/s13244-023-01541-3.

[79] K.G. van Leeuwen, M.J. Becks, D. Grob, F. de Lange, J.H.E. Rutten, S. Schalekamp, M.J.C.M. Rutten, B. van Ginneken, M. de Rooij, F.J.A. Meijer, AI-support for the detection of intracranial large vessel occlusions: One-year prospective evaluation, Heliyon 9 (2023) e19065, https://doi.org/10.1016/j.heliyon.2023.e19065.

[80] A. Bhandari, 2024. Revolutionizing Radiology With Artificial Intelligence. Cureus 16, e72646. doi: 10.7759/cureus.72646.

[81] B. Kim, S. Romeijn, M. van Buchem, M.H.R. Mehrizi, W. Grootjans, A holistic approach to implementing artificial intelligence in radiology, Insights Imaging 15 (2024) 22, https://doi.org/10.1186/s13244-023-01586-4.

[82] J. Schneider, I. Smalley. CPU vs. GPU for machine learning. https://www.ibm.com/think/topics/cpu-vs-gpu-machine-learning Published January 15, 2025. Accessed March 07, 2025.

[83] B. Schmidt, A. Hildebrandt, From GPUs to AI and quantum: three waves of acceleration in bioinformatics, Drug Discov. Today 29 (2024) 103990, https://doi.org/10.1016/j.drudis.2024.103990.

[84] A. Kirimtat, O. Krejcar, GPU-based parallel processing techniques for enhanced brain magnetic resonance imaging analysis: a review of recent advances, Sensors (Basel) 24 (2024) 1591, https://doi.org/10.3390/s24051591.

[85] R. Najjar, Redefining radiology: a review of artificial intelligence integration in medical imaging, Diagnostics (Basel) 13 (2023) 2760, https://doi.org/10.3390/diagnostics13172760.

[86] D. Garisto, How cutting-edge computer chips are speeding up the AI revolution, Nature 630 (2024) 544–546, https://doi.org/10.1038/d41586-024-01544-0.

[87] F. van der Vlist, A. Helmond, F. Ferrari, Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence, Big Data Soc. 11 (2024), https://doi.org/10.1177/20539517241232630.

[88] P. Ahrweiler, Towards Culture-Sensitive, Responsive, and Participatory AI, in: P. Ahrweiler (ed), Participatory Artificial Intelligence in Public Social Services. Artificial Intelligence, Simulation and Society, Springer, 2025, pp. 277-306. doi: 10.1007/978-3-031-71678-2_13.

[89] I. Golgeci, P. Ritala, A. Arslan, B. McKenna, I. Ali, Confronting and alleviating AI resistance in the workplace: an integrative review and a process framework, Hum. Resour. Manag. Rev. 35 (2025) 101075, https://doi.org/10.1016/j.hrmr.2024.101075.

[90] R.G. Cooper, Why AI projects fail: lessons from new product development, IEEE Eng. Manag. Rev. 52 (2024) 15–21, https://doi.org/10.1109/EMR.2024.3419268.

[91] N. Stogiannos, E. Georgiadou, N. Rarri, C. Malamateniou, Ethical AI: a qualitative study exploring ethical challenges and solutions on the use of AI in medical imaging, EJR AI 1 (2025) 100006, https://doi.org/10.1016/j.ejrai.2025.100006.

[92] R. Cheraghi, H. Ebrahimi, N. Kheibar, M.H. Sahebihagh, Reasons for resistance to change in nursing: an integrative review, BMC Nurs. 22 (2023) 310, https://doi.org/10.1186/s12912-023-01460-0.

[93] L. Ranasinghe, F.J. Dor, P. Herbert, Turning the oil tanker: a novel approach to shifting perspectives in medical practice, Adv. Med. Educ. Pract. 10 (2019) 507–511, https://doi.org/10.2147/amep.s197570.

[94] N. Agarwal, A. Moehring, P. Rajpurkar, T. Salz. Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology. https://www.nber.org/system/files/working_papers/w31422/w31422.pdf Updated March 2024. Accessed April 11, 2025.

[95] B. Koçak, A. Ponsiglione, A. Stanzione, C. Bluethgen, J. Santinha, L. Ugga, M. Huisman, M.E. Klontzas, R. Cannella, R. Cuocolo, Bias in artificial intelligence for medical imaging: fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects, Diagn. Interv. Radiol. 31 (2025) 75–88, https://doi.org/10.4274/dir.2024.242854.

[96] The Royal College of Radiologists. Overcoming Barriers to AI Implementation in Imaging. https://www.rcr.ac.uk/media/05mp1eda/overcoming_barriers_to_ai_implementation_in_imaging_v3.pdf Published June 2023. Accessed April 11, 2025.

[97] B. Kocak, A.H. Yardimci, S. Yuzkan, A. Keles, O. Altun, E. Bulut, O.N. Bayrak, A. A. Okumus, Transparency in artificial intelligence research: a systematic review of availability items related to open science in radiology and nuclear medicine, Acad. Radiol. 30 (2023) 2254–2266, https://doi.org/10.1016/j.acra.2022.11.030.

[98] F. Gunzer, M. Jantscher, E.M. Hassler, T. Kau, G. Reishofer, Reproducibility of artificial intelligence models in computed tomography of the head: a quantitative analysis, Insights Imaging 13 (2022) 173, https://doi.org/10.1186/s13244-022-01311-7.

[99] B.S. Kelly, C. Judge, S.M. Bollard, S.M. Clifford, G.M. Healy, A. Aziz, P. Mathur, S. Islam, K.W. Yeom, A. Lawlor, R.P. Killeen, Radiology artificial intelligence: a systematic review and evaluation of methods (RAISE), Eur. Radiol. 32 (2022) 7998–8007, https://doi.org/10.1007/s00330-022-08784-6.

[100] R. Han, J.N. Acosta, Z. Shakeri, J.P.A. Ioannidis, E. Topol, P. Rajpurkar, Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review, Lancet Digit. Health 6 (2024) e367–e373, https://doi.org/10.1016/S2589-7500(24)00047-5.

[101] M. Huisman, B. van Ginneken, H. Harvey, The emperor has few clothes: a realistic appraisal of current AI in radiology, Eur. Radiol. 34 (2024) 5873–5875, https://doi.org/10.1007/s00330-024-10664-0.

[102] B. Koçak, F. Köse, A. Keleş, A. Şendur, I. Meşe, M. Karagülle, Adherence to the Checklist for Artificial Intelligence in Medical Imaging (CLAIM): an umbrella review with a comprehensive two-level analysis, Diagn. Interv. Radiol. (2025), https://doi.org/10.4274/dir.2025.243182.

[103] E.P. Larsen, T. Hailu, L. Sheldon, A. Ginader, N. Bodo, D. Dewane, A.J. Degnan, J. Finley, R.W. Sze, Optimizing radiology reading room design: the eudaimonia radiology machine, J. Am. Coll. Radiol. 18 (2021) 108–120, https://doi.org/10.1016/j.jacr.2020.09.041.

[104] J.S.N. Tang, J.K.C. Lai, J. Bui, W. Wang, P. Simkin, D. Gai, J. Chan, D.M. Pascoe, S.B. Heinze, F. Gaillard, E. Lui, Impact of different artificial intelligence user interfaces on lung nodule and mass detection on chest radiographs, Radiol. Artif. Intell. 5 (2023) e220079, https://doi.org/10.1148/ryai.220079.

[105] M. Demirer, S. Candemir, M.T. Bigelow, S.M. Yu, V. Gupta, L.M. Prevedello, R. D. White, J.S. Yu, R. Grimmer, M. Wels, A. Wimmer, A.H. Halabi, A. Ihsani, T. P. O'Donnell, B.S. Erdal, A User interface for optimizing radiologist engagement in image data curation for artificial intelligence, Radiol. Artif. Intell. 1 (2019) e180095, https://doi.org/10.1148/ryai.2019180095.

[106] H. Liu, N. Ding, X. Li, Y. Chen, H. Sun, Y. Huang, C. Liu, P. Ye, Z. Jin, H. Bao, H. Xue, Artificial intelligence and radiologist burnout, JAMA Netw. Open 7 (2024) e2448714, https://doi.org/10.1001/jamanetworkopen.2024.48714.

[107] P. Rajpurkar, C. O'Connell, A. Schechter, N. Asnani, J. Li, A. Kiani, R.L. Ball, M. Mendelson, G. Maartens, D.J. van Hoving, R. Griesel, A.Y. Ng, T.H. Boyles, M. P. Lungren, CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV, NPJ Digit. Med. 3 (2020) 115, https://doi.org/10.1038/s41746-020-00322-2.

[108] M. Sujan, R. Pool, P. Salmon, Eight human factors and ergonomics principles for healthcare artificial intelligence, BMJ Health Care Inform. 29 (2022) e100516, https://doi.org/10.1136/bmjhci-2021-100516.

[109] H.J. Shin, K. Han, L. Ryu, E.K. Kim, The impact of artificial intelligence on the reading times of radiologists for chest radiographs, NPJ Digital Med. 6 (2023) 82, https://doi.org/10.1038/s41746-023-00829-4.

[110] K. Wenderott, J. Krups, J.A. Luetkens, N. Gambashidze, M. Weigl, Prospective effects of an artificial intelligence-based computer-aided detection system for prostate imaging on routine workflow and radiologists' outcomes, Eur. J. Radiol. 170 (2024) 111252, https://doi.org/10.1016/j.ejrad.2023.111252.

[111] M.H. Bernstein, M.K. Atalay, E.H. Dibble, A.W.P. Maxwell, A.R. Karam, S. Agarwal, R.C. Ward, T.T. Healey, G.L. Baird, Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography, Eur. Radiol. 33 (2023) 8263–8269, https://doi.org/10.1007/s00330-023-09747-1.

[112] L. Lind Plesner, F.C. Müller, M.W. Brejnebøl, L.C. Laustrup, F. Rasmussen, O. W. Nielsen, M. Boesen, M. Brun Andersen, Commercially available chest radiograph AI tools for detecting airspace disease, pneumothorax, and pleural effusion, Radiology 308 (2023) e231236, https://doi.org/10.1148/radiol.231236.

[113] S. Gaube, H. Suresh, M. Raue, A. Merritt, S.J. Berkowitz, E. Lermer, J.F. Coughlin, J.V. Guttag, E. Colak, M. Ghassemi, Do as AI say: susceptibility in deployment of clinical decision-aids, NPJ Digit. Med. 4 (2021) 31, https://doi.org/10.1038/s41746-021-00385-9.

[114] F. Yu, A. Moehring, O. Banerjee, T. Salz, N. Agarwal, P. Rajpurkar, Heterogeneity and predictors of the effects of AI assistance on radiologists, Nat. Med. 30 (2024) 837–849, https://doi.org/10.1038/s41591-024-02850-w.

[115] L. Rettenberger, M. Schilling, S. Elser, M. Bohland, M. Reischl, Self-supervised learning for annotation efficient biomedical image segmentation, I.E.E.E. Trans. Biomed. Eng. 70 (2023) 2519–2528, https://doi.org/10.1109/tbme.2023.3252889.

[116] P. Singh, R. Chukkapalli, S. Chaudhari, L. Chen, M. Chen, J. Pan, C. Smuda, J. Cirrone, Shifting to machine supervision: annotation-efficient semi and self-supervised learning for automatic medical image segmentation and classification, Sci. Rep. 14 (2024) 10820, https://doi.org/10.1038/s41598-024-61822-9.

[117] E.M. Weisberg, L.C. Chu, B.D. Nguyen, P. Tran, E.K. Fishman, Is AI the Ultimate QA? J. Digit. Imaging 35 (2022) 534–537, https://doi.org/10.1007/s10278-022-00598-8.

[118] The Royal College of Radiologists. Cancer and diagnostic waiting times for November 2024. https://www.rcr.ac.uk/news-policy/latest-updates/cancer-and-diagnostic-waiting-times-for-november-2024/ Published January 9, 2025. Accessed March 17, 2025.

[119] J. Lam Shin Cheung, A. Ali, M. Abdalla, B. Fine. U"AI" Testing: User Interface and Usability Testing of a Chest X-ray AI Tool in a Simulated Real-World Workflow. Can. Assoc. Radiol. J. 74 (2023) 314-325. doi: 10.1177/08465371221131200.

[120] F. Rosa, D. Buccicardi, A. Romano, F. Borda, M.C. D'Auria, A. Gastaldo, Artificial intelligence and pelvic fracture diagnosis on X-rays: a preliminary study on performance, workflow integration and radiologists' feedback assessment in a spoke emergency hospital, Eur. J. Radiol. Open 11 (2023) 100504, https://doi.org/10.1016/j.ejro.2023.100504.

[121] K. Palaniappan, E.Y.T. Lin, S. Vogel, J.C.W. Lim, Gaps in the global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector and key recommendations, Healthcare (Basel) 12 (2024) 1730, https://doi.org/10.3390/healthcare12171730.

[122] M. McKee, O.J. Wouters, The challenges of regulating artificial intelligence in healthcare comment on "clinical decision support and new regulatory frameworks for medical devices: are we ready for it? - a viewpoint paper", Int. J. Health Policy Manag. 12 (2023) 7261, https://doi.org/10.34172/ijhpm.2022.7261.

[123] H. van Kolfschooten, J. van Oirschot, The EU Artificial Intelligence Act (2024): Implications for healthcare, Health Policy 149 (2024) 105152, https://doi.org/10.1016/j.healthpol.2024.105152.

[124] M. Huisman, E. Ranschaert, W. Parker, D. Mastrodicasa, M. Koci, D. Pinto de Santos, F. Coppola, S. Morozov, M. Zins, C. Bohyn, U. Koç, J. Wu, S. Veean, D. Fleischmann, T. Leiner, M.J. Willemink, 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 (2021) 7058–7066, https://doi.org/10.1007/s00330-021-07781-5.

[125] N. Stogiannos, G. Walsh, B. Ohene-Botwe, K. McHugh, B. Potts, W. Tam, C. O'Sullivan, A.S. Quinsten, C. Gibson, R.G. Gorga, D. Sipos, E. Dybeli, M. Zanardo, C. Sá Dos Reis, N. Mekis, C. Buissink, A. England, C. Beardmore, A. Cunha, A. Goodall, J.S. John-Matthews, M. McEntee, Y. Kyratsis, C. Malamateniou, R-AI-diographers: a European survey on perceived impact of AI on professional identity, careers, and radiographers' roles, Insights Imaging 16 (2025) 43, https://doi.org/10.1186/s13244-025-01918-6.

[126] Health & Care Professions Council. Radiographers. Standards of Proficiency. https://www.hcpc-uk.org/globalassets/resources/standards/standards-of-proficiency—radiographers.pdf Published September 2023. Accessed March 08, 2025.

[127] R. van de Venter, E. Skelton, J. Matthew, N. Woznitza, G. Tarroni, S.P. Hirani, A. Kumar, R. Malik, C. Malamateniou, Artificial intelligence education for radiographers, an evaluation of a UK postgraduate educational intervention using participatory action research: a pilot study, Insights Imaging 14 (2023) 25, https://doi.org/10.1186/s13244-023-01372-2.

[128] G. Walsh, N. Stogiannos, R. van de Venter, C. Rainey, W. Tam, S. McFadden, J. P. McNulty, N. Mekis, S. Lewis, T. O'Regan, A. Kumar, M. Huisman, S. Bisdas, E. Kotter, D. Pinto Dos Santos, C. Sá Dos Reis, P. van Ooijen, A.P. Brady, C. Malamateniou, Responsible AI practice and AI education are central to AI implementation: a rapid review for all medical imaging professionals in Europe, BJR Open 5 (2023) 20230033, https://doi.org/10.1259/bjro.20230033.

[129] J.R. Wilson, L.M. Prevedello, C.D. Witiw, A.E. Flanders, E. Colak, Data liberation and crowdsourcing in medical research: the intersection of collective and artificial intelligence, Radiol. Artif. Intell. 6 (2024) e230006, https://doi.org/10.1148/ryai.230006.

[130] F. Busch, J.N. Kather, C. Johner, M. Moser, D. Truhn, L.C. Adams, K.K. Bressem, Navigating the European union artificial intelligence act for healthcare, NPJ Digit. Med. 7 (2024) 210, https://doi.org/10.1038/s41746-024-01213-6.

[131] G.S. Collins, K.G.M. Moons, P. Dhiman, R.D. Riley, A.L. Beam, B. Van Calster, M. Ghassemi, X. Liu, J.B. Reitsma, M. van Smeden, A.L. Boulesteix, J. C. Camaradou, L.A. Celi, S. Denaxas, A.K. Denniston, B. Glocker, R.M. Golub, H. Harvey, G. Heinze, M.M. Hoffman, A.P. Kengne, E. Lam, N. Lee, E.W. Loder, L. Maier-Hein, B.A. Mateen, M.D. McCradden, L. Oakden-Rayner, J. Ordish, R. Parnell, S. Rose, K. Singh, L. Wynants, P. Logullo, TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods, BMJ 385 (2024) e078378, https://doi.org/10.1136/bmj-2023-078378.

[132] B. Vasey, M. Nagendran, B. Campbell, D.A. Clifton, G.S. Collins, S. Denaxas, A. K. Denniston, L. Faes, B. Geerts, M. Ibrahim, X. Liu, B.A. Mateen, P. Mathur, M. D. McCradden, L. Morgan, J. Ordish, C. Rogers, S. Saria, D.S.W. Ting, P. Watkinson, W. Weber, P. Wheatstone, P. McCulloch, Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI, BMJ 377 (2022) e070904, https://doi.org/10.1136/bmj-2022-070904.

[133] H. Ibrahim, X. Liu, S.C. Rivera, D. Moher, A.W. Chan, M.R. Sydes, M.J. Calvert, A. K. Denniston, Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines, Trials 22 (2021) 11, https://doi.org/10.1186/s13063-020-04951-6.

[134] N. Stogiannos, C. Gillan, H. Precht, C.S.D. Reis, A. Kumar, T. O'Regan, V. Ellis, A. Barnes, R. Meades, M. Pogose, J. Greggio, E. Scurr, S. Kumar, G. King, D. Rosewarne, C. Jones, K.G. van Leeuwen, E. Hyde, C. Beardmore, J.G. Alliende, S. El-Farra, S. Papathanasiou, J. Beger, J. Nash, P. van Ooijen, C. Zelenyanszki, B. Koch, K.A. Langmack, R. Tucker, V. Goh, T. Turmezei, G. Lip, C.C. Reyes-Aldasoro, E. Alonso, G. Dean, S.P. Hirani, S. Torre, T.N. Akudjedu, B. Ohene-Botwe, R. Khine, C. O'Sullivan, Y. Kyratsis, M. McEntee, P. Wheatstone, Y. Thackray, J. Cairns, D. Jerome, A. Scarsbrook, C. Malamateniou, A multidisciplinary team and multiagency approach for AI implementation: a commentary for medical imaging and radiotherapy key stakeholders, J. Med Imaging Radiat. Sci. 55 (2024) 101717, https://doi.org/10.1016/j.jmir.2024.101717.

[135] J. Westenberger, K. Schuler, D. Schlegel, Failure of AI projects: understanding the critical factors, Procedia Comput. Sci. 196 (2022) 69–76, https://doi.org/10.1016/j.procs.2021.11.074.

[136] J. Ryseff, B.F. De Bruhl, S.J. Newberry. The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed. https://www.rand.org/pubs/research_reports/RRA2680-1.html Published August 13, 2024. Accessed March 09, 2025.

[137] D. Schlegel, K. Schuler, J. Westenberger, Failure factors of AI projects: results from expert interviews, IJSPM 11 (2023) 25–40, https://doi.org/10.12821/ijispm110302.