

Central Lancashire Online Knowledge (CLOK)

Title	Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice
Type	Article
URL	https://knowledge.lancashire.ac.uk/id/eprint/57125/
DOI	https://doi.org/10.1136/bmjdhai-2025-000073
Date	2025
Citation	Thomson, Rachel M orcid iconORCID: 0000-0002-3060-939X, Perdomo-Lampignano, Jesus, Fisher, Euan, Wati-tsayo, Cindy Karelle, Jeyakumar, Gowsikan, Duncan, Sean and Lowe, David J (2025) Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice. <i>BMJ Digital Health & AI</i> , 1 (1). bmjdhai-2025.
Creators	Thomson, Rachel M, Perdomo-Lampignano, Jesus, Fisher, Euan, Wati-tsayo, Cindy Karelle, Jeyakumar, Gowsikan, Duncan, Sean and Lowe, David J

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.1136/bmjdhai-2025-000073>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLOK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Central Lancashire Online Knowledge (CLOK)

Title	Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice
Type	Article
URL	https://clock.uclan.ac.uk/id/eprint/57125/
DOI	https://doi.org/10.1136/bmjdhai-2025-000073
Date	2025
Citation	Thomson, Rachel M orcid iconORCID: 0000-0002-3060-939X, Perdomo-Lampignano, Jesus, Fisher, Euan, Wati, Cindy, Jeyakumar, Gowsikan, Duncan, Sean and Lowe, David J (2025) Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice. BMJ Digital Health and AI, 1 (1). bmjdhai-2025-000073. ISSN 3049-575X
Creators	Thomson, Rachel M, Perdomo-Lampignano, Jesus, Fisher, Euan, Wati, Cindy, Jeyakumar, Gowsikan, Duncan, Sean and Lowe, David J

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.1136/bmjdhai-2025-000073>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLOK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice

Rachel M Thomson ^{1,2} Jesus Perdomo-Lampignano,^{1,2} Euan Fisher,³ Cindy Wati,⁴ Gowsikan Jeyakumar,³ Sean Duncan,^{1,3} David J Lowe^{1,3}

To cite: Thomson RM, Perdomo-Lampignano J, Fisher E, *et al.* Evaluating the environmental sustainability of AI in radiology: a systematic review of current practice. *BMJ Digital Health and AI* 2025;1:e000073. doi:10.1136/bmjdhai-2025-000073

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjdhai-2025-000073>).

Received 9 April 2025
Accepted 29 August 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY. Published by BMJ Group.

¹University of Glasgow, Glasgow, UK

²NHS Education for Scotland, Edinburgh, UK

³NHS Greater Glasgow & Clyde, Glasgow, UK

⁴University of Central Lancashire, Preston, UK

Correspondence to

Dr Rachel M Thomson;
rachel.thomson@glasgow.ac.uk

ABSTRACT

Objective Data-heavy and energy-heavy artificial intelligence (AI) technologies are increasingly being applied in healthcare, particularly for clinical imaging, often without consideration of their environmental impacts. We aimed to assess current practice in considering and evaluating environmental sustainability (ES) impacts of AI-enabled clinical pathways in radiology.

Methods and analysis We searched MEDLINE and Embase on 5 November 2024 for quantitative clinical radiology studies which used a machine learning approach to aid in radiological diagnosis or intervention and discussed its ES impacts. We included peer-reviewed, English language studies published from 2015 onwards. Our primary outcome was any quantitative reporting of ES impacts, and our secondary outcome was any within-text qualitative discussion of ES impacts. For quantitative outcomes, we conducted synthesis without meta-analysis based on effect direction and size, with our secondary outcome synthesised narratively.

Results Of 4449 records screened, 18 met our inclusion criteria. Six (33.33%) reported quantitative ES outcomes and 15 (83.33%) included qualitative discussion of ES. When applied to the same tasks, algorithms designed to be 'lightweight' demonstrated from 2.19 to 17.15 times reduction in carbon emissions (median 7.81, 16 datapoints) and from 1.60 to 751.62 times reduction in energy consumption (median 3.22, 16 datapoints) compared with state-of-the-art alternatives, while maintaining similar or improved clinical performance. No quantitative studies compared ES outcomes for an AI-enabled pathway versus standard-of-care, and 75.00% of studies reporting only on our secondary outcome included just a single sentence on sustainability.

Conclusion Despite increasing concern about the climate impacts of AI, environmental outcomes are rarely considered within evaluations of AI technologies in clinical radiology. However, there are approaches available with smaller carbon footprints. To meet their stated aims on sustainability, funders and governance bodies should consider how to promote integration of environmental impact assessment into AI health research and evaluation.

PROSPERO registration number CRD42024601818.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Artificial intelligence (AI) technologies are increasingly used in healthcare, with higher energy and data demands than traditional digital tools; however, environmental sustainability is rarely considered in evaluations of AI-enhanced clinical pathways.

WHAT THIS STUDY ADDS

⇒ This is the first systematic review to examine how environmental sustainability is assessed in healthcare evaluations of AI, focusing on clinical radiology.
⇒ Very few studies quantitatively evaluated environmental impacts, and these were limited to AI tools specifically designed to be more resource-efficient.
⇒ These studies all reported a reduced environmental impact compared with state-of-the-art competitors—on average, an eight times reduction in carbon emissions and three times reduction in energy consumption—while achieving similar or improved clinical performance.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Environmental impacts of AI are largely overlooked in current healthcare evaluations, although some tools do have smaller carbon footprints.
⇒ Incorporating sustainability metrics into AI assessments is essential to support responsible innovation and avoid unintended contributions to the climate crisis.

INTRODUCTION

The integration of artificial intelligence (AI) technologies into healthcare has increased rapidly in the last decade with the emergence of powerful deep learning algorithms that can reliably perform complex tasks, particularly in image recognition.¹ While the theory required to facilitate these advances has existed for decades, recent progress has been driven by increasing availability and decreasing cost of two key commodities: 'big data' to train AI models, and sufficient computational power to process and store such data.² These data-intensive and energy-intensive AI

algorithms are now being deployed throughout clinical pathways, working on tasks ranging from segmentation of tumours and nodules to clinical risk prediction.³ Radiology as a specialty has been a particularly early adopter of AI, producing around a third of all published clinical AI research.⁴

The myriad potential applications mean AI is often touted as a panacea for productivity and efficiency, but the environmental impacts of these technologies are rarely considered in clinical practice.^{5 6} This is despite there being considerable cause for concern⁷: global AI energy demand is projected to exceed that of a country the size of Belgium by 2026,⁸ and the data centres required for model development have significant and increasing impacts on material environments, water consumption, mineral extraction and e-waste disposal.⁶ However, it is worth noting that any efficiencies resulting from AI do have potential to reduce carbon emissions in clinical pathways compared with usual practice, which must be adequately taken into account when considering their overall environmental impact.⁹

Given that healthcare systems will be responsible for managing the increased care burden resulting from the health impacts of the climate crisis, while being major emitters of greenhouse gases themselves, there is a strong argument that clinicians should have a role in the climate response.¹⁰ One significant way in which clinicians can influence healthcare-related emissions is through involvement in decision-making on sustainable healthcare provision and procurement.^{10 11} However, to facilitate the inclusion of sustainability as a factor in decision-making on adoption and deployment of novel technologies like AI in healthcare, their environmental impacts (both positive and negative) must be transparently measured and reported on.¹²

A recent scoping review suggests environmental assessment remains poorly integrated into most forms of health technology assessment (HTA),¹³ potentially due to the associated technical and resource challenges.¹⁴ Despite environmental sustainability being a known issue in AI and data-driven research, progress here is particularly slow, with little engagement on the issues of environmental sustainability at all.^{6 7} To our knowledge, there are no existing systematic reviews on this topic specific to AI, and no available guidance on how to select, measure or evaluate sustainable AI software.⁵ We therefore aimed to assess current practice in considering and evaluating the potential or actual environmental sustainability impacts of AI-enabled clinical pathways, focusing on radiology to capitalise on the relative maturity of AI research within this field, and its increasing prioritisation of sustainability.^{4 15 16}

METHODS

We conducted a methodological systematic review, pre-registering our protocol with PROSPERO (CRD42024601818). Reporting follows the Preferred

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Synthesis Without Meta-analysis (SWiM) in Systematic Reviews guidance.^{17 18} The minimal protocol deviations are reported in online supplemental appendix A.

Literature search

Searches of peer-reviewed literature were conducted in MEDLINE and Embase on 5 November 2024. To ensure inclusion of studies that were broadly representative of current practice, we designed our search strategy to be in line with that of Kelly *et al* in their 2022 systematic review, which aimed to synthesise the body of evidence on AI use in radiology at that point in time.³ Searches included terms for AI combined with the AND operator for terms relevant to radiology AND sustainability (full search strategy in online supplemental appendix B).

Inclusion criteria

The relevant population of interest was clinical radiological papers (ie, describing the use of imaging techniques to diagnose, treat or monitor diseases and injuries), which were hospital-based, focused on human participants and published from 2015 onwards. Radiology in this context includes all radiographic, computed tomography (CT), magnetic resonance (MR), ultrasound or nuclear medicine/molecular or hybrid imaging techniques. There were no restrictions on the type of participants in the included studies. The intervention of interest was any machine learning tool used to aid care of a patient's radiological diagnosis or intervention, which aimed to complete a segmentation, identification, classification or prediction task using computer vision techniques. A comparator group was not required for inclusion.

Our primary outcome group of interest was any quantification or measurement of variables related to environmental sustainability. Specific measures prespecified in the protocol to be included in our primary outcome group were carbon dioxide emissions, other measures of energy usage, water usage, impacts on the material environment, mineral extraction or usage, and waste generation and disposal. Our secondary outcome group of interest was any acknowledgement of or reference to potential environmental sustainability impacts of the intervention within the text of the study or the supplementary materials. This included any qualitative discussion of the following terms or topics, even if not formally quantified: environmental sustainability, climate, carbon dioxide emissions, energy, water and minerals usage, impacts on material environment, and waste generation and disposal.

We included all quantitative, peer-reviewed study types that met our population criteria, including conference abstracts, to ensure inclusion of the most up-to-date AI research.

Exclusion criteria

We excluded non-English papers and grey literature, that is, literature which was not formally published in books/journals or peer-reviewed (including preprints). In keeping with Kelly *et al*,³ we also excluded non-human, laboratory and phantom-based studies; studies of functional MRI; studies solely for use in radiation therapy; radiomics studies and studies focusing on texture analysis or identification of imaging biomarkers; connectomics and quality assessment studies; image processing and registration studies; and image quality studies.

Study selection

References were de-duplicated in EndNote 21 and imported to Covidence for screening. All title/abstracts and full texts were independently screened by RMT and a second reviewer (JPL, EF, GJ or CW), with conflicts resolved by consensus and/or discussion with a third reviewer (SD). Reviewer topic/specialty expertise included radiology, public health, oncology and digital innovation in healthcare, with reviewers ranging in stage from medical student (n=1) to senior specialty trainee with ≥ 3 years of specialty training (n=5).

Relevant commentary articles and reviews captured by the searches were excluded at title and abstract stage but tagged within Covidence to facilitate hand searching of reference lists. Reference lists of included studies were also screened for additional studies.

Data extraction and synthesis

Data were extracted in Excel by RMT and independently checked by a second reviewer (JPL). The data extraction template is available on GitHub.¹⁹ As this was a methodological systematic review interested primarily in the extent of reporting on the topic, it was relatively agnostic to the AI methods used or the study conduct. Given this fact, and the fact that the included studies were anticipated to be highly heterogeneous,³ no formal risk of bias assessment was completed.

For our primary outcome, we followed SWiM guidance with vote counting based on direction of effect and summarising of effect estimates.²⁰ We summarised effect estimates as they were reported in the majority of studies, as unstandardised relative differences (comparator/intervention); it was not possible to reliably convert the absolute difference to a standardised, comparable metric. For our secondary outcome (qualitative discussion of environmental sustainability), we summarised findings narratively. We intended to stratify findings by study characteristics of interest where possible, for example, retrospective versus prospective, study size, study subspecialty, AI task, but were unable to do so due to the small number of quantitative datapoints.

Where relevant, we documented the economic perspective taken in the evaluation (eg, healthcare provider/sector, health system, societal²¹) as well as the time horizon selected to model health and sustainability costs and benefits. Finally, we assessed which of four proposed

approaches to HTA environmental impact assessment the study took: information conduit (where an HTA simply republishes existing environmental data without further assessment), integrated evaluation (where clinical, financial and environmental information is combined into a single quantitative analysis), parallel evaluation (where environmental data are analysed and presented separately from established health economic analyses) or environment-focused evaluation (where technologies which do not have a direct health benefit are evaluated solely on their environmental benefits).¹⁴

RESULTS

Study characteristics

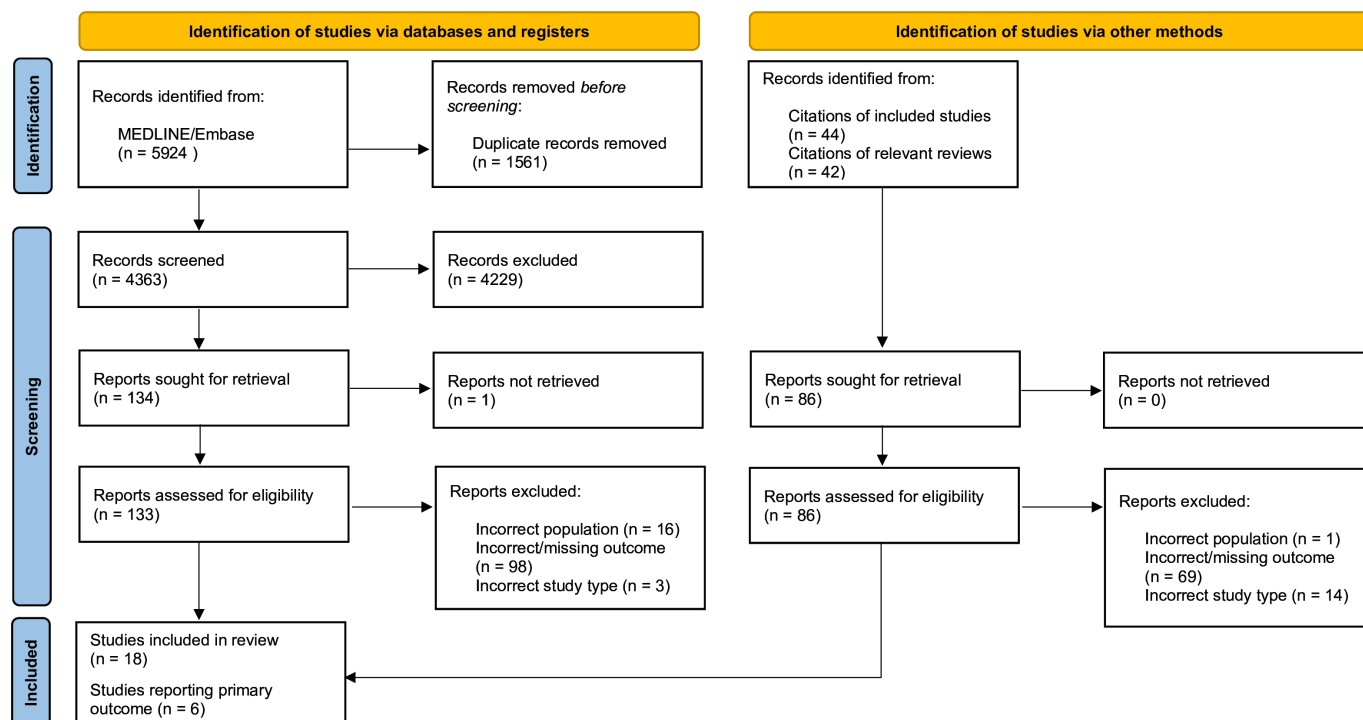
Of 4449 citations screened, 18 were eligible for inclusion (figure 1). Six studies^{22–27} reported our primary outcome of quantitative environmental sustainability measures (32 datapoints), and 15 studies^{22 25 26 28–39} included qualitative discussion of environmental sustainability (16 datapoints). Most of the studies were peer-reviewed journal articles (n=14, 77.78%) and the remainder were peer-reviewed conference abstracts. A list of studies excluded at full-text stage is included in online supplemental appendix C. Reviewer agreement was 97.22%, with only seven studies requiring consensus discussion at full-text stage; the kappa statistic on inter-rater reliability for full-text screening was 0.73, indicating substantial agreement.⁴⁰

The majority of included studies (n=12, 66.67%) were from high-income countries, with only two countries contributing more than one study (the USA n=7, China n=4). Most studies analysed retrospective secondary data (n=16, 88.89%), most commonly public data (n=7, 38.89%). Interventions were most often for cancer imaging (n=8, 44.44%), and the most imaged sites were head (n=6, 33.33%), prostate (n=5, 27.78%) and chest (n=2, 11.11%). MRI was the most common imaging modality (n=13, 72.22%), with ultrasound and X-ray being the next most common (n=2, 11.1%) and one study using CT.

The most common AI task conducted as part of the intervention was classification (n=7, 38.89%), followed by segmentation (n=5, 27.78%) and prediction (n=4, 22.22%). Most studies used supervised rather than unsupervised learning (n=14, 77.78%). A range of AI architectures was used, with the most common being Green Learning (n=5, 27.78%)⁴¹ and various forms of U-Net (eg, two-dimensional and three-dimensional variants) (n=8, 44.44%). Full details of included studies are shown in table 1; all extracted data are available on GitHub.¹⁹

Primary outcome: quantitative measures of sustainability

The six studies which quantitatively measured environmental sustainability reported on carbon emissions (CO₂eq, n=3, 16 datapoints) or other measures of energy consumption (eg, joules/watts, n=5, 16 datapoints), and all reported on algorithms *a priori* intended



Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

This work is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart.

to be more efficient or less computationally intensive than competitors (table 1). Two studies reported use of carbon tracking software packages^{42 43} to generate their outcomes; no others reported how emissions or energy consumption were calculated. One study reporting on energy consumption²³ could not be included in quantitative analysis, as findings were only reported in a graph; we attempted to contact the authors for the underlying data, but this was not successful.

We identified no studies reporting on any other sustainability measures of interest, for example, water/mineral usage, waste generation/disposal or impacts on material environments. The comparators in all six studies were other AI interventions applied to the same clinical task, with no studies comparing sustainability outcomes for an AI-enabled pathway versus standard-of-care. All quantitative studies took a healthcare provider perspective and conducted a parallel evaluation of environmental outcomes (ie, analysing and presenting environmental data alongside health analyses, rather than directly integrating this into health economic models or having an entirely separate environment-focused evaluation¹⁴). Half of the studies measured outcomes during algorithm training only, with the other half measuring outcomes during deployment only.

For all primary outcome datapoints, the AI intervention of interest had a beneficial effect direction (ie, was shown to produce fewer carbon emissions and consume less energy than the comparator models) when applied to the same clinical tasks, while maintaining similar or improved clinical performance (table 1). Intervention

algorithms generated from 2.19 to 17.15 times less carbon emissions (median 7.81) and from 1.60 to 751.62 times less energy consumption (median 3.22) (table 2). The heterogeneity of included measures precluded further quantitative synthesis, for example, meta-analysis or standardisation.

Secondary outcome: qualitative discussion of sustainability

The 15 studies that included qualitative discussion of sustainability discussed energy usage (n=8), environmental sustainability or climate change (n=4), and impacts on material environment (n=4). Studies which mentioned material environment impacts all focused on the use of gadolinium-based contrast agent (GBCA) and referred to its role as a wastewater pollutant. In most papers, particularly those that did not include quantitative measurements of sustainability, discussion of sustainability was very limited: 75% of these (n=9) included only a single sentence on sustainability, either mentioning the use of a lightweight algorithm designed to be 'green', or briefly noting environmental concerns related to energy consumption or pollutants. The extracted data are included in full in table 1.

The three studies that reported both our primary and secondary outcome^{22 25 26} included much more in-depth discussion of the environmental sustainability aspects of AI in radiology, with this being a clear focus of the paper. These studies discussed their own 'green' intentions with their models and referenced the wider clinical and policy context, making recommendations for both greater computational efficiency of algorithms to reduce

Table 1 Table of included studies

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Primary outcome: quantitative measurement of environmental sustainability							
Bohoran <i>et al</i> ⁴²	Aortic lumen segmentation to automate aortic distensibility calculation	UK	Retrospective analysis of data from four UK trials involving cardiac MRI (2009–2018). Adults aged 18–79 years, mix of healthy controls and those with type II diabetes, coronary artery dissection, hypertension. Sample size=376	BConvLSTM U-Net architecture used to design deliberately resource-efficient algorithm. Resource-efficient features included fewer filters in convolutional layers, fewer steps involving BConvLSTM than standard, and only one densely packed convolutional block in the final encoding step	1. SOTA algorithm for the same purpose, less resource efficient 2. Unpruned algorithm for the same purpose Both trained using the same multicentre, multivendor, multidisease dataset following similar training and hyperparameter tuning procedures	1. CO2eq emissions (g) 2. Energy spent (kWh) 3. Equivalent distance travelled by car (km) All measured during final model training using Carbontracker Python tool ⁴² generating 2.9 times less carbon emissions.	The lightweight study algorithm outperformed the SOTA comparators on all measures of segmentation accuracy while consuming 3.9 times less fuel and generating 2.9 times less carbon emissions.
Hu <i>et al</i> ²³	Brain lesion segmentation and diagnosis prediction	China	Retrospective analysis of brain MRI data obtained from a hospital; no further information on data source or population. Sample size not reported	Hybrid Pyramid U-Net architecture and fuzzy clustering used to create more efficient algorithm for brain image processing and brain disease diagnosis prediction	1. Convolutional Neural Network 2. Recurrent Neural Network 3. Fuzzy C-Means 4. Local Density Clustering Fuzzy C-Means 5. Adaptive Fuzzy C-Means All included in simulations for performance comparison; no detail on how trained/tuned	Total network energy consumption (J) per simulation time/round Unclear how measured. Outcome reported only in graph, and therefore unable to be extracted for inclusion in synthesised quantitative results (authors contacted for further detail but no response)	The study algorithm had slightly reduced energy consumption when compared with five comparators applied to the same task, while achieving the highest prediction and segmentation accuracy. Authors claim this shows improved energy efficiency of fuzzy clustering as an approach, but did not explicitly link this to sustainability.
Jyoti <i>et al</i> ²⁴	Automatic diagnosis of COVID-19 via chest X-ray classification	India	Retrospective analysis of subset of data from public chest X-ray datasets for COVID-19 (COVIDx and COVID-CXNet), collected 2013–2017. Sample size=7468	2D tunable Q-wavelet transform based on MCA used for initial stage of image decomposition, before pretrained ResNet50 and AlexNet models used to classify images into COVID-19 and non-COVID-19. The MCA-based model is designed to process and store data more efficiently than alternatives	Use of conventional digital CMOS-based computing rather than MCA-based approach for three tasks: 1. Image compression 128x128 2. Image compression 512x512 3. Full adder circuit	1. Energy consumption (nJ/pJ) 2. Power consumption (pJ/μW) Unclear how measured. Comparator estimates appear to be sourced from existing literature rather than rerunning the models	The MCA-based approach to image decomposition resulted in up to 752 times reduction in energy consumption and up to 300 times improvement in power consumption for specific tasks compared with conventional CMOS-based computing, but the authors did not explicitly link this to sustainability.
Serrador <i>et al</i> ²⁵	Vertebrae segmentation	Portugal and Italy	Retrospective analysis of data from the public VerSe challenge dataset: large-scale, multidetector, multisite CT spine data collected 2013–2017. Sample size=300	Two-step segmentation. Knowledge distillation methods were implemented by training a teacher network with performance similar to that found in the literature, and this knowledge was distilled to a shallower student network	Other SOTA DL-based approaches: 1. Several 3D U-Net 2. Spatial-configuration Net 3. Spine-Transformers 4. Cyclic SS-VS-VI	CO2 emissions during segmentation of all test scans in VerSe dataset Measured using CodeCarbon package ⁴³	The 'shallower' networks used achieved comparable segmentation performance to SOTA models and demonstrated up to 89.36% reduction in carbon emissions.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Siouras <i>et al</i> ²⁶	ACL tear classification	Greece	Retrospective analysis of two public datasets (MRNet (2001–2012) and KneeMRI (2006–2014)) and one validation-only hospital dataset (2018). Sample size=1631	Generated highly discriminative novelty score by leveraging aleatoric semantic uncertainty as modelled in class scores outputted by a YOLOv5-nano object detection model. Second module of the pipeline is the MINIROCKET timeseries classification model to determine presence or absence of an ACL tear	Two SOTA approaches to ACL classification: 1. Jeon <i>et al</i> 2021 2. Dunnhofer <i>et al</i> 2022	1. CO2eq emissions (g) 2. Energy spent (kWh) 3. Equivalent distance travelled by car (km) Authors state calculated during final training, but method of calculation not reported	The novel method outperformed classification of SOTA alternatives on accuracy and sensitivity, while generating at least 2.1 times less carbon emissions and consuming at least 2.6 times less energy.
Xiong <i>et al</i> ²⁷	Brain tumour segmentation	China	Retrospective analysis of data from the BraTS19 and BraTS20 public datasets, produced for multimodal brain tumour MRI segmentation challenges in 2019 and 2020. Sample size not reported	FPGA-based inference accelerator used to speed up quantisation process, quantising and retraining a 3D U-Net neural network for batch normalisation layers to reduce parameter size and computational complexity	Compared with a traditional GPU-based approach: 1. Before quantisation 2. After quantisation	Energy consumption (W) Unclear how this was calculated, but appears to correspond to individual uses of the method, rather than the training period	FPGA-based processing provided more time efficient and accurate segmentation results, with energy efficiency 11.22 times and 82.33 times more efficient than traditional CPU and GPU-based processing. However, the authors did not explicitly link this to environmental sustainability.
Secondary outcome: qualitative discussion of environmental sustainability							
Abreu <i>et al</i> ²⁸	Prostate lesion detection	USA	Retrospective analysis of the public PI-CAI challenge dataset, an annotated multicentre dataset of bpMRIs to validate algorithm performance at detection and diagnosis of grade group ≥2 prostate lesions. Sample size=1300	Green Learning paradigm used for feature extraction, using IPHOP-II method to decompose input and feed to the classifier, which is trained to discern between prostate cancer and non-prostate cancer	None	Authors report use of Green Learning architecture, designed to be more environmentally sustainable: "We used the Green Learning paradigm for feature extraction, which offers a lightweight model size and explainable feature extraction process"	Authors note that their choice of model is lightweight compared with alternative approaches, which is linked specifically to 'green' aims in the selection of the Green Learning algorithm.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Bohoran <i>et al</i> ²²	Aortic lumen segmentation to automate aortic distensibility calculation	UK	As above	As above	As above	Authors explicitly link their algorithm design to sustainability repeatedly throughout the paper, for example, "The network... is resource-efficient to help promote environmentally friendly and more inclusive DL research and practices"; "By making energy usage and greenhouse gas emissions explicit, the presented work aligns with efforts to keep DL's energy requirements and carbon cost in check"	Authors argue for energy usage and emissions to be considered key evaluation criteria in DL research, given their increasing use and significant contribution to carbon footprints: "The value of DL models should be judged by the amount of intelligence they provide per joule".
Cacciamani <i>et al</i> ²³	Prostate lesion classification	USA	Retrospective analysis of database of men undergoing prostate MRI and biopsy from 2017 to 2023 from a separate study by the same author group (IPB# HS-13-00663). Sample size=259	Green Learning framework used for feature extraction. Data-driven automatic, unsupervised process (IPHoP-II) used to generate spatial-spectral representation of region of interest, before feature selection module is used to filter out noise and best features fed to classifier to classify according to prostate imaging reporting and data system	None	Authors report use of Green Learning architecture, designed to be more environmentally sustainable: "A novel Green Learning framework with a lightweight model and explainable feature extraction process was used"	Authors note that their choice of model is lightweight compared with alternative approaches, which is linked specifically to 'green' aims in the selection of the Green Learning algorithm.
Guldogan <i>et al</i> ³⁰	Breast lesion classification	Turkey	Retrospective analysis of data collected from women scheduled for ultrasound-guided core needle biopsy across three breast imaging centres, which were part of the same Turkish institution from August 2021 to May 2022. Sample size=530	Koios DS for Breast System used to provide probability of malignancy (BI-RADS) based on orthogonal views of lesion	None	Authors discuss implications of successful AI for sustainability in healthcare: "In conclusion, AI-based DS system performance in our study was comparable with that of breast radiologists in predicting the final assessment of breast masses. The implementation into the clinical workflow may lead to a reduction of unnecessary biopsies and short-term follow-ups, which, in turn, can contribute to sustainability in healthcare practices"	Authors link a potential reduction in unnecessary biopsies and follow-up appointments with sustainable healthcare.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Haase <i>et al</i> ⁸¹	Prediction of full-dose GBGA images from non-contrast/low-contrast brain MRI	Germany	Prospective analysis of data from adults planned for contrast-enhanced MRI brain at single centre from August to October 2021, who received three 3D T1w scans (non-contrast, 20% standard dose, full dose). Sample size=213	Three previously published DL approaches were used to predict the full dose scan, all of which were U-Net based	U-Net algorithm performances compared with each other.	Authors discuss the environmental impacts of overuse of gadolinium-based contrast: "Reducing gadolinium-based contrast agents in brain MRI is a research topic with several motivators... [including] ... lowering the impact on the environment by gadolinium-containing wastewater"	Authors state that a motivation for studies like theirs is to reduce the dose of gadolinium-containing contrast required for patients, which has patient benefits but also wider environmental benefits as gadolinium is excreted into wastewater and has harmful environmental impacts.
Kaneko <i>et al</i> ⁸²	Prostate lesion detection	USA	Retrospective analysis of database of men undergoing prostate MRI and biopsy from 2017 to 2023 from a separate study by the same author group (IRB# HS-13-00663). Sample size=259	Green Learning framework used for feature extraction and ROI classification, specifically the IPHO-P-II process. Independent modules used for peripheral and transition zones	None	Authors report use of Green Learning architecture, designed to be more environmentally sustainable: "For feature extraction and ROI classification, we adopt the Green Learning paradigm, and specifically the IPHO-P-II feature extraction process"	Authors use the Green Learning model, which is designed to be lightweight and linked explicitly to green aims through its name. However, authors themselves do not comment on model complexity or sustainability.
Kaneko <i>et al</i> ⁸³	Automated prostate segmentation	USA	Retrospective analysis of prostate MRI and biopsy from 2017 to 2023 from a separate study by the same author group (IRB# HS-13-00663). Sample size=119	Novel two-stage automatic Green Learning-based model. First stage segments prostate gland and second stage zooms in to delineate the transition zone and peripheral zone. Both stages share a lightweight feed-forward encoder-decoder Green Learning system	None	Authors report use of Green Learning architecture, designed to be more environmentally sustainable: "A lightweight feed-forward encoder-decoder model based on Green Learning can precisely segment the whole prostate"	Authors note that their choice of model is lightweight compared with alternative approaches, which is linked specifically to 'green' aims in the selection of the Green Learning algorithm.
Lee <i>et al</i> ⁸⁴	Ultrasound image reconstruction and inference	South Korea	Prospective analysis of data from volunteers, who received abdominal ultrasound; no information provided on recruitment. Most images normal. Sample size=25	DL edge-computing method for mobile device-based ultrasound scanners that perform image reconstruction and inference for object classification. Four architectures used: AlexNet, ShallowNet, MobileNet and Xception	All four algorithms trained on RSC-generated dataset versus standard ultrasound dataset	Authors discuss that the rollout of cloud-based AI is limited by a number of factors including energy consumption: "in a cloud-based ultrasonic artificial intelligence system, the use of POCUS is limited due to information security, network integrity and significant energy consumption"	Authors note that cloud-based systems are associated with significant energy consumption, which is particularly problematic in resource-poor settings. They propose their algorithms as an alternative that can run without requiring internet access.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Li <i>et al</i> ³⁵	COVID-19 detection on chest X-ray	China	Retrospective analysis of data from five public data repositories of COVID-19 images, creating the COVIDx dataset. Sample size=16 689	FedFocus, a FL platform which allows clients to collaboratively train models without exposing their source data. ResNet18 architecture used for training model in each client	1. Central training (not possible in real world due to data privacy restrictions) 2. Local training (client takes training set allocated to it for individual training) 3. FedAvg (leading method in FL to minimise training loss)	Authors discuss reducing energy consumption as a goal throughout, for example, “Current FL-based COVID-19 detection methods tend to deploy the most popular methods FedAvg or FedSGD, and optimize multiobjectives including delay, energy consumption, and privacy”; “Compared with other deeper residual network models, ResNet18 has fewer parameters effectively relieve the pressure of network bandwidth and energy consumption”	Authors selected an AI architecture that is more lightweight with fewer parameters and explicitly tie this to potential reduction in energy consumption of their final model.
Magouliantits <i>et al</i> ³⁶	Prostate cancer segmentation	USA	Retrospective analysis of public dataset of bpMRI data generated for the PI-CAI prostate segmentation challenge. Sample size=1000	Two-stage method which first generates voxel-wise predictions to deduce prostate cancer probability of each voxel, and second derives an anomaly score heatmap providing initial indication of suspicious looking areas with high probability of cancer	Other SOTA DL approaches: 1. SwinTransformer 2. nnDetection 3. nnU-Net 4. U-Net	Authors report use of Green Learning architecture, designed to be more environmentally sustainable: “It [the approach] adopts the recently introduced Green Learning paradigm, which offers a small model size and low complexity”; “this stresses one of the key benefits which is to offer lightweight and environment-friendly solutions”	Authors note that their choice of model, named ‘Green Learning’ to emphasise its reduced environmental impact, is lightweight compared with alternative approaches using DL, and they explicitly link this design choice to environmental benefits.
Pasumarthi <i>et al</i> ³⁷	Prediction of full-dose GBCA images from non-contrast/low-contrast brain MRI	USA	Retrospective analysis of data from three sites of patients undergoing specific enhanced scanning protocol on brain MRI. Sample size=640	U-Net based model to predict contrast-enhanced images from corresponding precontrast and low-contrast images. Quantitative tumour segmentation performed on cases with abnormal enhancements	1. Low-dose MRI 2. Full-dose MRI (compared segmentation outcome)	Authors discuss the environmental impacts of overuse of GBCA: “Environmental sustainability concerns are also being raised as gadolinium is an emerging water pollutant”	Authors link a reduction in use of GBCA with a reduction in water pollution.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Serrador <i>et al</i> ²⁵	Vertebrae segmentation	Portugal and Italy	As above	As above	As above	Authors highlight the large environmental impact of existing models and note the favourability of their approach, for example, "The size and complexity of these DL models contribute to a heightened environmental footprint"; "four method] ... contributed) to reduced environmental impact as outlined in the European Commission's Ethics Guidelines for Trustworthy AI"	Authors aimed to create models that were explicitly more lightweight and had smaller environmental footprints than comparative algorithms, linking this goal to ethical guidance on building trustworthy AI.
Siouras <i>et al</i> ²⁶	ACL tear classification	Greece	As above	As above	As above	Both environmental sustainability and energy usage were explicitly highlighted throughout the article, for example, "concerns over the environmental footprint of DL research and its impact on accelerating global climate change have been growing"; "[this] research paves the way for future advancements in the DL-powered MRI-based injury detection, emphasizing environmental sustainability alongside technological innovation"; "previous papers disregarded the compute demand by the DL model. Yet, this is a significant matter since these computations necessitate mind-boggling amounts of generated power for fueling them"	Authors selected a model structure that is less computationally intensive and specifically link this to environmental sustainability, highlighting wider discussion within DL literature on green/climate goals. The authors also highlight the current lack of consideration of computational demands of DL models and express the view that this is an oversight given the high energy consumption associated with complex models.
Wang <i>et al</i> ³⁸	Prediction of full-dose GBGA images from non-contrast brain MRI	China	Retrospective analysis of data from patients undergoing brain MRI at two hospitals in China from 2019 to 2021. Unclear if recruited specifically for this study. Sample size=221	3D U-Net-like generator producing 3D isotropic full-contrast T2Flair images from 2D anisotropic non-contrast T2Flair image stacks	1. 3D Res U-Net 2. Bayes U-Net 3. Convnets	Authors discuss the environmental impacts of overuse of GBGA: "In addition, gadolinium has been reported as an emerging water pollutant and may harm the environment"	Authors link a reduction in use of GBGA with a reduction in water pollution.

Continued

Table 1 Continued

Study	AI task	Setting	Methods	Intervention	Comparator(s)	Outcome(s)	Summary of findings
Xie <i>et al</i> ³⁹	Prediction of full-dose GBCA images from non-contrast brain MRI	USA	Retrospective analysis of the public BraTS20 dataset, produced for multimodal brain tumour segmentation challenge in 2020. Sample size=369	Two sequential networks: 1. A retina U-Net trained to derive semantic features from non-contrast MR images to represent tumour regions 2. A synthesis module trained to generate synthetic contrast-enhanced MR images from semantic feature maps and non-contrast images	None	Authors discuss the environmental impacts of overuse of GBCA: "Although GBCAs are generally thought to be safe, various health and environmental concerns have been raised recently about their use in MR imaging"; "Debates on toxicity of gadolinium deposition in the brain, breast, and body, as well as environmental issues, have been raised by most recent studies"	Authors designed a machine learning approach that would reduce the amount of GBCA required in MRI studies, and link the desire to do so with environmental as well as health concerns about contrast use.

ACL, anterior cruciate ligament; AI, artificial intelligence; BI-RADS, Breast Imaging Reporting and Data System; CMOS, complementary metal oxide semiconductor circuits; CO₂eq, carbon dioxide equivalent; CPU, central processing unit; 2D, two-dimensional; 3D, three-dimensional; DL, deep learning; DS, decision support; FL, Federated Learning; FPGA, field programmable gate array; GBCA, gadolinium-based contrast agent; GPU, graphics processing unit; IPHO-P-II, integrated Phage-Host Prediction tool; MCA, memristive crossbar array; MR, magnetic resonance; PI-CAL, prostate imaging-cancer artificial intelligence; POCUS, point-of-care ultrasound; ROI, region of interest; RSC, reverse scan conversion; SOTA, State-of-the-art; T1w, T1-weighted.

the carbon footprint of AI in radiology, and increased reporting of sustainability outcomes as standard alongside other common outcomes. Interestingly, however, the three other studies that reported quantitative measures (all measures of energy consumption) did not explicitly link this to environmental sustainability at any point in their papers.^{23 24 27}

DISCUSSION

Across a 9-year period when AI-related publications in clinical radiology were rapidly increasing from ~100 to almost 4000/year,⁴⁴ our methodological systematic review found very few studies of AI interventions in radiology discussing environmental sustainability (n=15) and even fewer (n=6) quantitatively measuring sustainability outcomes. Quantitative outcomes reported included carbon emissions and energy consumption (eg, kilowatt hours or joules), with studies finding that their algorithms were able to provide similar or better clinical performance while on average emitting eight times less carbon and consuming three times less energy. Half of these studies did not explicitly link their measurement of such outcomes to environmental sustainability, and all studies were of algorithms that were specifically designed to be lightweight and less computationally intensive than competitors. Other than in the three papers which had a deliberate focus on environmental sustainability,^{22 25 26} qualitative discussion of the topic was very superficial, with most studies including only a single sentence on the topic. Interestingly, a common topic raised in four papers was the potential for AI technologies to reduce the amount of GBCA required for scans, with this having the environmental co-benefit of reducing gadolinium pollution of wastewater.^{31 37-39}

As noted in our introduction, it has been suggested that environmental assessment is poorly integrated into HTAs.^{13 14} Our findings suggest that this is certainly the case for HTAs focusing specifically on evaluating AI technologies, at least within clinical radiology. In addition, despite the availability of existing frameworks—such as life cycle assessment⁴⁵ or integrated assessment¹⁴—the small number of studies reporting quantitative outcomes showed little consistency in what was reported. In several studies, a lack of reporting made it challenging to identify how or when these outcomes had been measured (table 1). This may suggest a requirement for effective guidance and frameworks that are more tailored to AI technology evaluations in healthcare.⁴⁶ The most comprehensive studies^{22 25} leveraged software tools such as Carbontracker or CodeCarbon,^{42 43} which allowed them to report more standardised outcomes, for example, CO₂eq emissions, or energy spent in kWh. More widespread awareness and adoption of these tools, which are designed to be simple to integrate into existing software pipelines and scalable, could also be beneficial in standardising outputs of environmental assessments within HTAs of AI healthcare technology.²⁶ However, this would

Table 2 Summary of relative differences (ratio)

Outcome	Mean	Median	Minimum	Maximum	IQR	Count (n)
Carbon emissions	8.39	7.81	2.19	17.15	7.07	16
Energy use	69.82	3.22	1.60	751.62	5.56	16

require upskilling of the clinical and technical workforce to ensure that integrating such software is seen as both feasible and valuable.

Of interest, many studies of algorithms designed for a similar purpose as those in our included studies (ie, to be lightweight) were excluded at full-text stage because the study authors did not link this intention to sustainability; instead, this was linked to goals such as promotion of mobile use of AI,⁴⁷ or improved use in resource-constrained settings.⁴⁸ This suggests that while there is clearly cost pressure to produce higher efficiency algorithms with lower computational demand, and while movement in this direction also has benefits for environmental sustainability, a desire for environmental sustainability as defined in this review is not currently a key driver of this. However, it demonstrates that if improved computational efficiency is seen as providing a competitive edge, then developers will focus on achieving this; perhaps a sign that, if incentives were to be altered in relation to sustainability, there could be a shift in the focus of developers and deployers.^{22 49}

A shift in incentives is potentially within the influence of decision-makers, as well as regulators, clinical guideline developers and research funders, many of whom have outlined in their own strategies a clear commitment to incorporating environmental sustainability assessment into their decision-making processes in future.^{50 51} Radiology as a specialty has also expressed a growing interest in AI's carbon footprint in recent years.^{5 15 16} Clear, structured guidance on how to report these outcomes, alongside a much firmer expectation from these bodies that they *should* be reported, could ideally lead to an increase in clinical radiology AI studies that transparently measure and report on the environmental performance of algorithms, allowing them to be more comprehensively judged according to 'the amount of intelligence they provide per joule'.²²

As the first on this topic, we believe our study has some important strengths. We closely followed a preregistered systematic review protocol and adhered to gold standard Cochrane guidance in our conduct, analysis and presentation of results.^{17 18 20} We built on an existing tested search strategy³ and believe we are the first to try to quantify the extent to which environmental sustainability is being measured in AI radiology literature to date. However, there are some important limitations to acknowledge. First, modifying the search strategy to find studies that considered sustainability was challenging. We were required to add terms for sustainability to keep the search outputs manageable, and while we allowed for these terms to be included anywhere that was indexed, it

is possible we may have missed some studies that briefly mention sustainability. However, as we note above, studies falling into this category did not contribute much to our synthesis, and we also performed citation searching to increase our yield, identifying two additional studies (figure 1). Second, as noted in the methods we have not formally measured study quality, so could not comment on how this may have influenced findings. However, given the consistency of the quantitative findings and small number of datapoints, it is unlikely this would have changed our conclusions. Finally, our ability to standardise effect sizes was limited by the uncommon and inconsistent measures used within the quantitative studies, meaning that we report only relative differences of the included measures.

Our findings have implications for both research and policy (table 3). We have identified a clear research gap for studies that compare AI-enabled clinical pathways to standard-of-care or care as usual, rather than just state-of-the-art alternative AI approaches; without this, it is impossible to adequately judge the potential environmental costs and benefits brought by the AI technologies. In terms of further evidence synthesis, given the pace of development (with a majority of our papers publishing from 2023 onwards), a living systematic review could be a useful methodology to explore this topic further in future,⁵² particularly with the recent rapid increase in papers using generative AI including large language models.⁵³ In addition, it would be of interest to apply our methods to other fields where AI technologies are beginning to be used more widely but where there is perhaps more awareness of environmental sustainability as a topic, for example, public health, to see whether this influences findings.⁵⁴ From a policy perspective, we have identified a clear need for a standardised methodology and reporting framework for environmental outcome assessment specific to AI, as evaluations are sparse and inconsistent in their methods and reporting. It is clear that reporting of these outcomes is not happening adequately at present, and it may be that policymakers must incentivise this in some way, for example, by ensuring there is some competitive advantage to reporting carbon emissions and energy consumptions of models when developers seek regulatory approval.²² In addition, given that many governments and healthcare systems have specific policies in place to mitigate against climate change, they risk directly undermining their efforts to do so if they pursue widespread AI deployment without adequate mitigation of environmental impact.

In conclusion, despite rising concerns about the potential impacts of AI technologies on the environment, measurement of sustainability is poorly integrated into current evaluations of AI in clinical radiology, except where reduced

Table 3 Summary of key recommendations

Recommendation	Target group(s)
Computational complexity and energy efficiency of AI algorithms should be more explicitly taken into account when procuring, training and testing these technologies in clinical settings	Researchers; clinicians; healthcare managers
Evaluations of AI deployment in clinical radiology must better integrate measurement and reporting of environmental sustainability, particularly when comparing AI-augmented pathways to standard-of-care	Researchers; clinicians
Development of effective guidance and/or frameworks for environmental evaluation of AI technology in healthcare would be beneficial, and should include recommendations on preferred metrics for reporting to improve comparability across studies	Regulators; clinical guideline developers; researchers; clinicians
AI developers seeking regulatory approval or access to clinical data for testing/deployment should be better incentivised to consider, measure and report environmental sustainability outcomes	Regulators; clinical guideline developers; funding bodies; clinicians
To facilitate increased reporting and ensure outcomes are more comparable across studies, there should be more widespread awareness and adoption of software tools designed to measure environmental outcomes of algorithms, for example, Carbontracker, CodeCarbon ^{42 43}	AI developers; researchers; clinicians
AI, artificial intelligence.	

computational demand was an explicit selling point of the algorithm. No studies compared AI with standard-of-care, making it impossible to ascertain whether AI-augmented pathways (particularly those focused on prevention) have potential to reduce future carbon costs associated with healthcare utilisation. Uncritically continuing the widespread rollout of AI in healthcare without evaluating its potential environmental impacts risks worsening the climate crisis, with major implications for global health. Researchers and clinicians should aim, where possible, to quantify the environmental impacts of the technologies they are evaluating, establishing the costs and benefits for the climate as well as for health. To meet their stated aims on sustainability, funders and governance bodies should consider how to promote better integration of environmental impact assessment into AI health research and evaluation.

Acknowledgements The authors gratefully acknowledge Ms Valerie Wells for her assistance in accessing Covidence during the search process.

Contributors RMT drafted the study protocol, with critical feedback provided by DJL and SD. RMT conducted literature searches and led screening, and JPL, EF, GJ and CW contributed to the screening process and selection of included studies. RMT extracted data, which were checked by JPL. RMT completed all data analysis, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors critically reviewed and approved the manuscript. RMT is the guarantor and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. The AI tool Elicit.com was used prior to writing the review protocol, alongside traditional scoping searches, to help identify any relevant existing systematic reviews and primary studies. AI was not used in the conduct of the review itself nor in the drafting of the manuscript text. Prior to submission to this journal, ChatGPT was used to assist in generating the short Key Messages summary from a longer prewritten summary, which had been drafted according to another journal's requirements.

Funding RMT and JPL are funded by NHS Clinical AI Fellowships (NHS Education for Scotland). DJL is funded by the Living Laboratory for Precision Medicine (UK Research and Innovation Strength in Places Fund, #107140) and has previously received grant support for studies from Qure.ai, US2.ai and Nanox.

Disclaimer The funders played no active role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; or decision to submit the manuscript for publication.

Competing interests All authors have completed the ICMJE uniform disclosure form at <http://www.icmje.org/disclosure-of-interest/>. RMT and JPL are funded by

NHS Clinical AI Fellowships (NHS Education for Scotland). DJL is funded by The Living Laboratory for Precision Medicine (UK Research and Innovation Strength in Places Fund #107140) and has previously received grant support for studies from Qure.ai, US2.ai and Nanox. All other authors declare no competing interests.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was not required for this project, as it was a systematic review of published research with no primary data collection.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The study protocol is published on PROSPERO: <https://www.crd.york.ac.uk/PROSPERO/view/CRD42024601818>. Data extracted from included studies (including data used for all analyses) and analytic code are available on GitHub: <https://github.com/rachelmthomson/thomson-AI-sustainability-sr>.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See <https://creativecommons.org/licenses/by/4.0/>.

ORCID iD

Rachel M Thomson <http://orcid.org/0000-0002-3060-939X>

REFERENCES

- Hosny A, Parmar C, Quackenbush J, *et al*. Artificial intelligence in radiology. *Nat Rev Cancer* 2018;18:500–10.
- Najjar R. Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. *Diagnostics (Basel)* 2023;13:2760.
- Kelly BS, Judge C, Bollard SM, *et al*. Radiology artificial intelligence: a systematic review and evaluation of methods (RAISE). *Eur Radiol* 2022;32:7998–8007.
- Stewart JE, Rybicki FJ, Dwivedi G. Medical specialties involved in artificial intelligence research: is there a leader. *TMJ* 2020;2:20–7.

- 5 Doo FX, Vosschenrich J, Cook TS, *et al.* Environmental Sustainability and AI in Radiology: A Double-Edged Sword. *Radiology* 2024;310:e232030.
- 6 Samuel G, Lucassen AM. The environmental sustainability of data-driven health research: A scoping review. *Digit Health* 2022;8:2055207622111297.
- 7 van Wynsberghe A. Sustainable AI: AI for sustainability and the sustainability of AI. *AI Ethics* 2021;1:213–8.
- 8 Ren S, Wierman A. The uneven distribution of AI's environmental impacts. *Harv Bus Rev* 2024.
- 9 Vinuesa R, Azizpour H, Leite I, *et al.* The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat Commun* 2020;11:233.
- 10 Braithwaite J, Pichumani A, Crowley P. Tackling climate change: the pivotal role of clinicians. *BMJ* 2023;382:e076963.
- 11 Wilburn S, Jharia I, Prabhakaran P. Sustainable procurement in healthcare. In: *Climate change and the health sector*. 2021: 183–91.
- 12 Dhar P. The carbon impact of artificial intelligence. *Nat Mach Intell* 2020;2:423–5.
- 13 Guirado-Fuentes C, Abt-Sacks A, Trujillo-Martin MDM, *et al.* Main Challenges of Incorporating Environmental Impacts in the Economic Evaluation of Health Technology Assessment: A Scoping Review. *Int J Environ Res Public Health* 2023;20:4949.
- 14 Toolan M, Walpole S, Shah K, *et al.* Environmental impact assessment in health technology assessment: principles, approaches, and challenges. *Int J Technol Assess Health Care* 2023;39:e13.
- 15 Rockall AG, Allen B, Brown MJ, *et al.* Sustainability in radiology: position paper and call to action from ACR, AOSR, ASR, CAR, CIR, ESR, ESRNM, ISR, IS3R, RANZCR, and RSNA. *Eur Radiol* 2025.
- 16 Indian Radiological and Imaging Association. Summary of the Proceedings of International Radiology Forum 2025: Green Radiology and Sustainability in Medical Imaging. *Indian J Radiol Imaging* 2025;35:209–12.
- 17 Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
- 18 Campbell M, McKenzie JE, Sowden A, *et al.* Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* 2020;368:l6890.
- 19 Data from: Thomson AI sustainability systematic review. Github; 2025. Available: <https://doi.org/10.5281/zenodo.16761455>
- 20 McKenzie JE, Brennan SE. Synthesizing and presenting findings using other methods. In: *Cochrane handbook for systematic reviews of interventions*. 2019: 321–47.
- 21 Sittimart M, Rattanavipapong W, Mirelman AJ, *et al.* An overview of the perspectives used in health economic evaluations. *Cost Eff Resour Alloc* 2024;22:41.
- 22 Bohoran TA, Parke KS, Graham-Brown MPM, *et al.* Resource efficient aortic distensibility calculation by end to end spatiotemporal learning of aortic lumen from multicentre multivendor multidisease CMR images. *Sci Rep* 2023;13:21794.
- 23 Hu M, Zhong Y, Xie S, *et al.* Fuzzy System Based Medical Image Processing for Brain Disease Prediction. *Front Neurosci* 2021;15:714318.
- 24 Jyoti K, Sushma S, Yadav S, *et al.* Automatic diagnosis of COVID-19 with MCA-inspired TQWT-based classification of chest X-ray images. *Comput Biol Med* 2023;152:106331.
- 25 Serrador L, Villani FP, Moccia S, *et al.* Knowledge distillation on individual vertebrae segmentation exploiting 3D U-Net. *Comput Med Imaging Graph* 2024;113:102350.
- 26 Siouras A, Moustakidis S, Chalatsis G, *et al.* Economical hybrid novelty detection leveraging global aleatoric semantic uncertainty for enhanced MRI-based ACL tear diagnosis. *Comput Med Imaging Graph* 2024;117:102424.
- 27 Xiong S, Wu G, Fan X, *et al.* MRI-based brain tumor segmentation using FPGA-accelerated neural network. *BMC Bioinformatics* 2021;22:421.
- 28 Abreu AL, Cacciamani G, Kaneko M, *et al.* MP09-06 ASSESSMENT OF A NOVEL BPMRI-BASED MACHINE LEARNING FRAMEWORK TO AUTOMATE THE DETECTION OF CLINICALLY SIGNIFICANT PROSTATE CANCER USING THE PI-CAI (PROSTATE IMAGING: CANCER AI) CHALLENGE DATASET. *J Urol* 2023;209.
- 29 Cacciamani G e., Kaneko M, Yang Y, *et al.* MP55-18 A NOVEL MACHINE LEARNING FRAMEWORK TO AUTOMATED CHARACTERIZE PROSTATE IMAGING REPORTING AND DATA SYSTEM (PIRADS) ON MRI. *J Urol* 2023;209.
- 30 Guldogan N, Taskin F, Icten GE, *et al.* Artificial Intelligence in BI-RADS Categorization of Breast Lesions on Ultrasound: Can We Omit Excessive Follow-ups and Biopsies? *Acad Radiol* 2024;31:2194–202.
- 31 Haase R, Pinetz T, Kobler E, *et al.* Artificial T1-Weighted Postcontrast Brain MRI: A Deep Learning Method for Contrast Signal Extraction. *Invest Radiol* 2025;60:105–13.
- 32 Kaneko M, Cacciamani GE, Magoulanis V, *et al.* MP55-20 A NOVEL MACHINE LEARNING FRAMEWORK FOR AUTOMATED DETECTION OF PROSTATE CANCER LESIONS CONFIRMED ON MRI-INFORMED TARGET BIOPSY. *J Urol* 2023;209:e771–2.
- 33 Kaneko M, Cacciamani GiE, Yang Y, *et al.* MP09-05 AUTOMATED PROSTATE GLAND AND PROSTATE ZONES SEGMENTATION USING A NOVEL MRI-BASED MACHINE LEARNING FRAMEWORK AND CREATION OF SOFTWARE INTERFACE FOR USERS ANNOTATION. *J Urol* 2023;209:e105–6.
- 34 Lee K, Kim M, Lim C, *et al.* Reverse Scan Conversion and Efficient Deep Learning Network Architecture for Ultrasound Imaging on a Mobile Device. *Sensors (Basel)* 2021;21:2629.
- 35 Li Z, Xu X, Cao X, *et al.* Integrated CNN and Federated Learning for COVID-19 Detection on Chest X-Ray Images. *IEEE/ACM Trans Comput Biol and Bioinf* 2024;21:835–45.
- 36 Magoulanis V, Yang J, Yang Y, *et al.* PCa-RadHop: A transparent and lightweight feed-forward method for clinically significant prostate cancer segmentation. *Comput Med Imaging Graph* 2024;116:102408.
- 37 Pasumathi S, Tamir JI, Christensen S, *et al.* A generic deep learning model for reduced gadolinium dose in contrast-enhanced brain MRI. *Magn Reson Med* 2021;86:1687–700.
- 38 Wang Y, Wu W, Yang Y, *et al.* Deep learning-based 3D MRI contrast-enhanced synthesis from a 2D noncontrast T2Flair sequence. *Med Phys* 2022;49:4478–93.
- 39 Xie H, Lei Y, Wang T, *et al.* Magnetic resonance imaging contrast enhancement synthesis using cascade networks with local supervision. *Med Phys Mex Symp Med Phys* 2022;49:3278–87.
- 40 Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977;33:159.
- 41 Kuo C-CJ, Madni AM. Green learning: Introduction, examples and outlook. *J Vis Commun Image Represent* 2023;90:103685.
- 42 Anthony LFW, Kanding B, Selvan R. Carbontracker: tracking and predicting the carbon footprint of training deep learning models. *arXiv [Preprint]* 2020.
- 43 Courty B, Schmidt V, Coutarel M, *et al.* mlco2/codecarbon: v2.2.7 (v2.2.7). Zenodo; 2023. Available: <https://doi.org/10.5281/zenodo.8181237>
- 44 Web of Science. Analyze results: artificial intelligence and radiology. 2025. Available: <https://www.webofscience.com/wos/allldb/analyze-results/605978c9-0cb3-495b-853a-08bd2e76c775-0158a5d45a>
- 45 Seifert C, Koep L, Wolf P, *et al.* Life cycle assessment as decision support tool for environmental management in hospitals: A literature review. *Health Care Manage Rev* 2021;46:12–24.
- 46 Ueda D, Walston SL, Fujita S, *et al.* Climate change and artificial intelligence in healthcare: Review and recommendations towards a sustainable future. *Diagn Interv Imaging* 2024;105:453–9.
- 47 Cao D, Zhang R, Zhang Y. MFLUnet: multi-scale fusion lightweight Unet for medical image segmentation. *Biomed Opt Express* 2024;15:5574–91.
- 48 Kumar G, Sharma N, Paul A. An extremely lightweight CNN model for the diagnosis of chest radiographs in resource-constrained environments. *Med Phys* 2023;50:7568–78.
- 49 Daniel-Watanabe L, Moore R, England N, *et al.* What is the carbon footprint of digital healthcare. 2024.
- 50 UK Research and Innovation (UKRI). UKRI environmental sustainability strategy 2020–2025. London, 2019.
- 51 National Institute for Health and Care Excellence (NICE). Environmental sustainability. 2024. Available: <https://www.nice.org.uk/about/who-we-are/sustainability>
- 52 Elliott JH, Synnot A, Turner T, *et al.* Living systematic review: 1. Introduction-the why, what, when, and how. *J Clin Epidemiol* 2017;91:23–30.
- 53 Akinci D'Antonoli T, Stanzione A, Bluethgen C, *et al.* Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions. *Diagn Interv Radiol* 2024;30:80–90.
- 54 Panteli D, Adib K, Buttigieg S, *et al.* Artificial intelligence in public health: promises, challenges, and an agenda for policy makers and public health institutions. *Lancet Public Health* 2025;10:e428–32.