Value assessment of artificial intelligence in medical imaging: a scoping review

Iben Fasterholdt1*, Mohammad Naghavi-Behzad2,3, Benjamin S. B. Rasmussen2,4,6, Tue Kjølhede1, Mette Maria Skjøth5, Malene Grubbe Hildebrandt1,2,3 and Kristian Kidholm1

Abstract

Background  Artificial intelligence (AI) is seen as one of the major disrupting forces in the future healthcare system. However, the assessment of the value of these new technologies is still unclear, and no agreed international health technology assessment-based guideline exists. This study provides an overview of the available literature in the value assessment of AI in the field of medical imaging.

Methods  We performed a systematic scoping review of published studies between January 2016 and September 2020 using 10 databases (Medline, Scopus, ProQuest, Google Scholar, and six related databases of grey literature). Information about the context (country, clinical area, and type of study) and mentioned domains with specific outcomes and items were extracted. An existing domain classification, from a European assessment framework, was used as a point of departure, and extracted data were grouped into domains and content analysis of data was performed covering predetermined themes.

Results  Seventy-nine studies were included out of 5890 identified articles. An additional seven studies were identified by searching reference lists, and the analysis was performed on 86 included studies. Eleven domains were identified: (1) health problem and current use of technology, (2) technology aspects, (3) safety assessment, (4) clinical effectiveness, (5) economics, (6) ethical analysis, (7) organisational aspects, (8) patients and social aspects, (9) legal aspects, (10) development of AI algorithm, performance metrics and validation, and (11) other aspects. The frequency of mentioning a domain varied from 20 to 78% within the included papers. Only 15/86 studies were actual assessments of AI technologies. The majority of data were statements from reviews or papers voicing future needs or challenges of AI research, i.e. not actual outcomes of evaluations.

Conclusions  This review regarding value assessment of AI in medical imaging yielded 86 studies including 11 identified domains. The domain classification based on European assessment framework proved useful and current analysis added one new domain. Included studies had a broad range of essential domains about addressing AI technologies highlighting the importance of domains related to legal and ethical aspects.

Keywords  Scoping review, Value assessment, Evaluation, Artificial intelligence, Medical imaging
Background

Artificial Intelligence (AI) includes various technologies based on advanced algorithms and learning systems. Different terms are used in connection with AI, such as machine learning, deep learning, and conventional neural networks [1]. Furthermore, there is no universally agreed-upon definition of AI, while it is suggested to define it as a system capable of interpreting and learning from data to produce a specific goal [2].

AI is seen as a digital transformation and could be one of the major disrupting tools in the future healthcare system [3]. Radiology and other imaging areas encompass a vast amount of manual image reviews and a steep increase in medical images has been observed in the last decade, which requires more interpretation times by the imaging specialist. This limiting factor could be reduced by incorporating computer-aided algorithms of machine learning into the clinical workflow [4]. Pattern recognition in medical images and AI technologies seems to be a good match [5], and this area is likely to be one of the first to benefit from AI, which provides high expectations. In the first quarter of 2019, funding in imaging AI companies exceeded 1.2 billion USD [6]. A consultancy company values the annual market for the top 10 AI-based healthcare solutions at 150 billion USD in 2026 [7]. Health care payers’ and providers’ expectations are to achieve cost savings, improve patient satisfaction, and optimise workforce resources [8].

Most of the published AI studies within medical imaging are retrospective with a technical focus, including reporting of clinical performance metrics, validation, or robustness of the model [9]. Evaluation of this phase is thoroughly described, e.g. in the guideline Checklist for Artificial Intelligence in Medical Imaging—CLAIM [10]. However, the lack of proven clinical utility, feasibility and effect on patient outcomes has been mentioned by several studies [11–14] as well as ethical, legal, economic, sharing of data, and implementation issues [15–17]. AI is a complex technology and implementing it in a complex healthcare system, critical to society, requires a broad framework.

Health technology assessment (HTA) provides a broad framework for valuing healthcare technologies and with several examples of being tailored for specific areas like telemedicine [18] and digital healthcare services [19]. Value is to be understood in a broad sense as referring to impact or effect in several different domains. HTA is a multidisciplinary process that summarizes information that has been collected in a systematic, transparent, unbiased, and robust manner. One example is the HTA-framework from EUnetHTA [20], where evaluation is performed from nine perspectives called “domains”: (1) the health problem and current use of technology; (2) description and technical characteristics of the new technology; (3) safety assessment; (4) clinical effectiveness; (5) economic evaluation; (6) ethical analysis; (7) organisational aspects; (8) patient and social aspects; and (9) legal aspects.

It is quite important to develop a holistic and tailored HTA tool for evaluating the value of AI to implement the correct AI technologies in the field of medical imaging. Therefore, we aimed to give a comprehensive overview of relevant identified domains in the literature when assessing the value of AI in medical imaging. This is the first step in creating an evidence-based assessment tool for valuing AI technology.

Methods

A scoping review aims to ‘map the key concepts underpinning a research area and the main sources and types of evidence available’ [21]. As such, scoping reviews typically address broad questions, potentially include a range of methodologies and do not undertake quality assessment. This contrasts with the focused question of a systematic review, which is answered from a relatively narrow range of quality-assessed studies. This scoping review was conducted based on the PRISMA guideline [22]. The study conducted at the Centre for Innovative Medical Technology at Odense University Hospital (Denmark), covering studies between January 2016 and September 2020, including five following stages.

Stage 1: Identifying the research issues

All published papers containing information about assessment of value of AI-technology in the field of medical imaging within public healthcare organisations, e.g. hospitals, dentists and universities, are considered eligible. Three types of studies are included to cover all aspects of AI in medical imaging: (1) Actual evaluation of the value of a specific AI technology, (2) Guidelines, statements, recommendations, white papers, evaluation models or HTA frameworks, and (3) Review articles and surveys voicing future needs/research/challenges. Hence, we reviewed available literature related to the assessment of the value of AI in medical imaging for a comprehensive understanding of which HTA domains or topics are used—or perceived relevant.

Stages 2 and 3: Identifying relevant studies and study selection

The search terms are summarised in Table 1. Clusters one and two are inspired by a previous review by Elhakim et al. [23] and cluster three by a search strategy used in a model for telemedicine [24]. The search strategy was verified by an experienced librarian from the University of Southern Denmark. Duplications were removed and
most of the article selection process was conducted in Covidence tool [25].

Literature searches was conducted during 17–18th September 2020 on following 10 databases: (1) Medline (Ovid), (2) Scopus, (3) ProQuest (includes EconLit), (4) Google Scholar, and six related grey literature and working paper resources, i.e. (5) International HTA Database [26], (6) OpenGrey [27], (7) National Institutes of Health [28], (8) National Health Services [29], (9) Folkehelseinstituttet [30], 10) Folkhälsomyndigheten [31]. A single backward snowballing was done on the 1–2 most central studies selected by each of the extractors to identify further studies. Developing the selection criteria was an iterative process, conducted by the entire research group through multiple consensus meetings.

Table 1  Search terms for performing a literature review

| Cluster 1: Technology       | MeSH terms  | Free text terms |
|-----------------------------|-------------|-----------------|
| Artificial intelligence     | Artificial intelligence |  |
| Machine learning            | Machine learning |  |
| Deep learning               | Deep learning |  |

| Cluster 2: Clinical area    | Diagnostic imaging | Diagnostic imaging |
|-----------------------------|---------------------|---------------------|
| Biomedical imaging          | Biomedical imaging |  |
| Medical imaging             | Medical imaging     |  |

| Cluster 3: Framework (and specific domains) | Delivery of health care | BROAD APPROACH: |
|--------------------------------------------|-------------------------|-----------------|
| Guidelines as topic                        | Checklist*               | FOCUS ON HEA & ORGANISATIONAL ISSUES: |
| Outcome and process assessment, Health care | Guide*                  | Health economic |
| Program evaluation                         | Framework*              | Cost effectiveness |
| Technology assessment, Biomedical Economics | HTA*                    | Cost utility analys* |
| Models, Organizational*                    | Assessment*             | Organizational |
| Ethics                                     | Evaluation Model*       |                  |

The search string was executed as [(Cluster 1) AND (cluster 2) AND (cluster 3)]

MeSH Medical Subject Headings

Table 2  Inclusion and exclusion criteria for selection of studies

| Inclusion criteria | 1. Included studies must cover domains/topics used—or mentioned as useful—for assessment of the value of AI in the area of medical imaging: |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------|
|                    | a. Actual evaluation or assessment of the value of a specific AI technology, i.e. AI interpretation, classification or pattern recognition of an image |
|                    | b. Guidelines, statements, recommendations, white papers, evaluation models or (HTA) frameworks described in the literature used for assessing the value of AI |
|                    | c. Reviews, surveys voicing future needs/research/challenges and future work when evaluating the value of AI |
|                    | 2. Studies must be published in English between January 1st 2016 and September 18th 2020 |
|                    | 3. Setting: Articles written within a hospital setting or public healthcare organisations are included, e.g., hospitals, dentists, universities, etc |

| Exclusion criteria | 4. Type of study or publication: All studies including grey literature, reports, and books but we exclude citations, patents, conference book, book of collected congress contributions, and conference abstracts |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------|
|                    | a. Because of many hits at the full-text stage, we later excluded: opinion, commentary, or viewpoint articles |
|                    | 5. Studies only reporting on clinical efficacy or performance metrics, validation studies, technical development of the prediction model or AI-model, i.e., studies focusing on reporting on, sensitivity, specificity, diagnostic accuracy, precision, AUC, software used, the robustness of the model, etc |
|                    | a. Technical interpretation of images, including image optimisation and enhancing methods for highlighting images |
|                    | 6. AI on non-clinical images like images of surgical equipment, plants and animals or not on images, e.g. AI on electronic health records or workflows, AI in drug development, surgical robotics, electrocardiogram, thermographic scans, brachytherapy treatment, or augmented reality visualization and VR |
(KK). The inclusion and exclusion criteria are shown in Table 2.

Stages 4 and 5: Charting the data and collating, summarising, and reporting the results
Initially, the required information on the included studies was extracted and summarised using an extraction template in Microsoft Office Excel software. Included studies were reviewed by two researchers independently. The extracted items were: Context (country, clinical area(s), type of study); mentioned domains in the study, including specific outcomes and items in each domain. The domain classification from EUnetHTA was used as a point of departure to group the extracted data into nine domains. If a topic did not fit these pre-existing domains, the extracted text was noted in the section of “other aspects” for later evaluation. Regarding the domains “clinical effectiveness” and “the health problem and current use of technology”, specific clinical outcomes are not part of the data extraction in this scoping review. We identified a new domain (development of AI algorithm, performance metrics and validation) but did not extract specific data as this domain has been thoroughly described already. For quality assurance purposes, the principal investigator (IF) had several bilateral consultations and group sessions with all extractors to solve any inconsistencies and challenges. Accordingly, a common understanding and agreement about which topics should go under which domains, how detailed or deep to extract data, arguments for excluded studies, was reached and helped align the raw data.

In the second phase, extracted data was analysed (by two researchers) and condensed for each domain. Qualitative content analysis was performed [32], covering four predetermined themes/questions: (1) Effects, outcomes, value or impacts mentioned—as well as future needs/challenges/topics relevant when evaluating the value of AI, (2) Specific outcome measures, (3) Frequency of using a given topic or outcome group, (4) Potential overlap with other domains. A summary for each domain was made as the last step of analysis.

Results
A study flow chart summarizes the process of literature retrieval (Fig. 1). In total, the literature search yielded 5890 papers, while 4292 papers remained after the removal of duplicates.

Based on titles and abstracts 166 papers were eligible for the full-text assessment, and a total of 79 papers [11, 12, 19, 33–108] fulfilled the inclusion criteria (see Additional file 1). Further, seven papers were included based on screening of reference lists in the most central articles [109–115]. Hence, 86 papers were included in the scoping review.

Search performed September 2020. Found 5890 records:
- Medline (Ovid): 1220. Scopus: 2203. ProQuest (includes EconLit): 2030. Google Scholar: 419. Grey literature and working paper resources: 18.

87 excluded*
1) Medical imaging or AI/ML/DL not mentioned or focus (n = 14)
1c) No mention of future needs/challenges/limitations in relation to evaluation of AI (n = 13)
2) Study not published in English (n=3)
4a) Exclude opinion, commentary, or viewpoint (N=11)
5) Clinical performance/efficacy/validation/technical model development study alone (n = 43)
6) AI on non-clinical images or not on images (N=3)

Fig. 1 PRISMA flow chart for selection of the studies. *The number in front of the list with exclusion reasons refers to the exclusion criteria in Table 2.
The characteristics of the 86 included studies are presented in Table 3. The most frequent clinical areas covered are radiology (17%), medical imaging (14%) and radiomics (13%). Most studies (n = 61) are voicing future needs or challenges when evaluating AI, while there were also studies containing an actual evaluation (n = 15) as well as guideline, statement or framework (n = 10). Most studies are published in 2019 and 2020. The most frequently mentioned or perceived relevant domains are the development of AI algorithm, performance metrics and validation (78%), followed by technology aspects (73%). Excluding “other aspects”, the variation between how often domains is mentioned is in the range of 20% to 78%.

In total, we identified eleven relevant domains about assessing the value of AI (two more than the EUneHTA
Table 4  Summary of extracted data for each specific domain

| Domain                                      | Topics or themes and outcome measures                                                                 |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------|
| (1) The health problem and current use of technology | No details extracted                                                                                |
| (2) Technology aspects                     | The main topic is interpretability in the sense that we need to avoid the “Black box problem” and the analysis done by the algorithm needs to be transparent to physicians/staff i.e., explainable AI |
|                                             | Furthermore, risk of bias, possibly causes discrimination issues and validation. The algorithm development method is highlighted, including data quality, the importance of annotation, external evaluation, and reference standards |
|                                             | Equipment and IT was a topic mentioning the clinical IT integration and infrastructure                  |
|                                             | **OUTCOME MEASURES:** interpretability, quality of scans, technical functioning/feasibility         |
| (3) Safety assessment                       | Safety of patients, potential challenges after implementation of AI to the healthcare system          |
|                                             | Reducing side effects and especially radiation dose, data security and protection                      |
|                                             | **OUTCOME MEASURES:** natural radiation exposure, using clinical knowledge support                  |
| (4) Clinical effectiveness, e.g. clinical outcomes* | No details extracted                                                                                |
| (5) Economics                               | The description of the savings and benefits are most often very general, e.g. improved cost-effectiveness |
|                                             | **OUTCOME MEASURES:** reduction in workload and time for staff, reduction in the number of biopsies and patients use of medication |
| (6) Ethical analysis                        | Privacy, consent, obligations, security, awareness of the use of patients’ data, and ownership of the data |
|                                             | Ethical approval and consider ethical issues of data, algorithms, trained models, and practice       |
|                                             | Understanding risks vs. benefits, shared/clear decision-making and transparency of results            |
|                                             | Big questions: “who owns data”, “can data and the algorithm be trusted” and “what is good clinical practice?” |
|                                             | **OUTCOME MEASURES:** validity of data, risks versus benefits, patient safety                         |
| (7) Organisational aspects                  | Benefits in the form of reductions in workflow and tasks related to imaging for the staff as a result of AI |
|                                             | The use of additional time related to implementation and training and the challenges related to ensuring acceptability |
|                                             | **OUTCOME MEASURES:** changes in time use for the health care professionals and patient, clinician acceptability measures |
| (8) Patients and social aspects             | Patients’ comfortability including easier imaging process and providing access to own data/report in a safe and secured platform |
|                                             | Better treatment outcome based on the improved clinical decision is the most discussed issue         |
|                                             | Patients’ satisfaction, as well as clinical benefits, could result in better acceptability of AI technology in the healthcare system |
|                                             | **OUTCOME MEASURES:** the time required for diagnosis, rating for overall satisfaction, help patients make more informed activity choices |
| (9) Legal aspects                           | Data security and privacy                                                                             |
|                                             | Responsibility for misdiagnosis                                                                       |
|                                             | **OUTCOME MEASURES:** regulatory approvals, consent from patients                                      |
| (10) Development of AI algorithm, performance metrics and validation | No details extracted                                                                                  |
| (11) Other aspects                          | Overpromising language in studies                                                                     |
|                                             | Offering the possibility of performing expensive and time-consuming screening programs in countries that otherwise cannot afford them |
|                                             | **OUTCOME MEASURES:** none identified                                                                  |

The full data analysis shows frequent overlap to other domains [see Additional file 2]

In addition, we extracted more specific data for eight domains including technology, safety, economics, ethics, organisational, patients and social, legal, and the other aspects. Regarding the remaining three domains (clinical effectiveness, the health problem and current use of technology, and development of AI algorithm, performance metrics and validation), we only noted whether information was
present or not. Table 4 contains summary and specific outcome measures (if applicable) through extracted information of the included domains. Further details can be found in Additional file 2.

**Clinical effects (domain 4)**
Although clinical outcomes are not part of the extraction of the included studies in the scoping review the clinical domain is evidently of great importance. Several outcome measures for clinical effectiveness studies have been suggested in a systematic scoping review of AI methods applied to adult patients who underwent any health/ medical intervention and reported therapeutic, preventive, or prognostic outcomes [116]. Based on 370 studies, this review found that AI was used primarily for the prediction/prognosis of more frequently reported outcomes, efficacy/effectiveness, and morbidity outcomes. Some examples of reasonably broad clinical or patient outcomes mentioned in the reviewed literature are fewer false positives in screening mammography, progression of disease and misdiagnosis, avoiding unnecessary stereotactic biopsies, improvement of treatment appropriateness (by physicians) and adherence (by patients) and the prevention of iatrogenic adverse events.

**Discussion**
**Main findings**
This scoping review regarding the value assessment of artificial intelligence in medical imaging yielded 86 included papers. Eleven domains were identified. The frequency of mentioning a domain varied from 20 to 78% within included studies. The studies were divided into three study types: studies voicing future needs or challenges when evaluating the value of AI, actual evaluations, and lastly, guidelines, statements, or frameworks. Out of 86 studies, only 15 were actual evaluations, and thus most data were based on statements and not actual outcomes of evaluations.

**Comparing findings to the literature**
The domain classification from the EUnetHTA framework proved very useful as extracted data used all the nine pre-existing domains and identified only one new domain (this is apart from a few issues filed in an “other domain”). However, the studies of value assessment of AI in the area of medical imaging includes a broad range of important domains in contrast to other studies. For example, an interview study in nine European countries investigated the information needs of hospital managers when deciding about investments in new treatments [117]). In that study, legal, social and ethical aspects were not deemed very important, which is in contrast with our findings. Further, in telemedicine, a scoping review of empirical studies that have applied the Model for Assessment of Telemedicine (MAST) shows that clinical, patient and economic effects are the most important areas [24]. Perhaps AI is unique in that all domains seem rather important—or perhaps in a more empirical setting and when more late evaluations become available this picture will change.

**Limitations and strengths**
Although the authors tried to provide a general overview regarding different aspects of value assessment for AI technology, there were some limitations. The terminology in the area of AI is still relatively immature. So although the focus was on medical imaging, we included articles in pathology and radiomics. Pathology and radiomics were not fully covered with our search terms, but this being a scoping review, we decided to keep these related studies and let the available data reflect in our analysis. Also, some of the included studies were narrative reviews that make it challenging to extract firm conclusions since there was no strong evidence to support mentioned results or claims. Another limitation was the considerable overlap between some domains where the included studies (and extractors) categorised data differently. This made it difficult to align the data analysis. We discussed overlaps in our joint meetings and selected the domain that best covered most of the included topics to handle these inconsistencies. For example, the topic of “explainable AI”/”black box” was initially extracted in both safety, technology, organisational and sometimes the other domain. Further, specific clinical outcomes are evidently of great importance when evaluating an AI technology. However, they were not part of the data extraction because most studies are very disease-specific and the number of included studies would have been unmanageable high in this review. This being a scoping review critical appraisal of the included studies was not done.

Regarding the strengths of this review, the broad coverage of areas of relevance for assessing the value of AI projects is in demand [17, 116, 118]. To the best of our knowledge, this is the first systematic scoping review about the value assessment of AI in field of medical imaging. Our search included grey literature which can be an important information source [119]. Furthermore, we have included the studies with different research methodologies to ensure the high coverage and a broad perspective in our data collection. The joint discussion sessions on topics with high overlap between the domains have likely improved the data analysis. High overlap was particularly observed in the
topics of patient's privacy (overlap between ethical, legal, and patient's domains) and the earlier mentioned black box topic. Furthermore, most studies are published recently (after 2019) which shows a great interest in this field and reflects the newest information on the topic. Drawing a framework based on data from recent publications can strengthen the current value assessment to be considered in future evaluations in this field.

Future directions
A pipeline for an overview of AI technologies evaluation is shown in Fig. 2. A majority of the published AI studies within medical imaging is in the retrospective phase of Fig. 2, i.e. having a technical focus. However, as our review shows, it is important to evaluate AI projects both clinically and concerning many other areas, i.e. the prospective phase in Fig. 2, to ensure that AI technologies with no effect or unintended effects are not uncritically implemented.

Results from this review was used as part of a project with the overall aim to develop a Model for ASsessing the value of AI in medical imaging (MAS-AI) cf. publication [120]. MAS-AI was developed in three phases. First, this literature review. Next, we interviewed leading researchers in AI in Denmark. The third phase consisted of two workshops where decision-makers, patient organizations, and researchers discussed crucial topics when evaluating AI. The multidisciplinary team revised the model between workshops according to comments. The HTA framework MAS-AI is to support the introduction of AI technologies into healthcare in medical imaging.

It is important to ensure uniform and valid decisions regarding the adoption of AI technology with a structured process and tool. The MAS-AI model can help support these decisions and provide greater transparency for all parties involved.

Conclusion
This scoping review regarding value assessment of artificial intelligence in medical imaging yielded 86 papers fulfilling the inclusion criteria, and eleven domains were identified: (1) the health problem and current use of technology, (2) technology aspects, (3) safety assessment, (4) clinical effectiveness, (5) economics, (6) ethical analysis, (7) organisational aspects, (8) patients and social aspects, (9) legal aspects, (10) development of AI algorithm, performance metrics and validation, and (11) other aspects. The domain classification from the EUnetHTA framework proved very useful and analysis identified only one new real domain: domain 10 (a few issues were included in an "other domain"). Studies include a broad range of essential domains when addressing AI technologies; in contrast to other areas, legal and ethical aspects are highlighted as important in this review.

Abbreviations
AI  Artificial intelligence
CLAIM  Checklist for Artificial Intelligence in Medical Imaging
HTA  Health technology assessment
Acknowledgements

The authors thank research librarian Mette Brandt Eriksen for help with narrowing down databases and providing technical help for searches and to trainee Laura Kuntz for conducting the searches and for honing the search criteria together with Iben Fasterholdt. Thanks to student assistant Jonathan Messerschmidt and Martin Granne Jensen for obtaining the full text of the included studies and Sara Eun Lendal for contribution to the linguistic content of the article and Valeria E. Rac for adding clarity.

Author contributions

IF, BR and KK conceived and designed the study. IF and TK did the literature search. IF, MNB, BR, TK, MWS and KK contributed to data collection and data analyses while IF did the data interpretation. IF, MNB, BR, TK, MMS, MGH and KK discussed the results and contributed to revise the final manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by an in-house fund at Odense University Hospital (Denmark) named ‘Konkurrencemidler’ in Danish. The funder had no role in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

The raw datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not relevant.

Competing interests

The authors declare that they have no competing interests.

Received: 22 July 2021 Accepted: 22 October 2022

Published online: 31 October 2022
79. Le Rest CC, Hustinx R. Are radiomics ready for clinical prime-time in PET/CT imaging? Q J Nucl Med Mol Imaging. 2019;63(4):347.
80. Lei W, Wang H, Gu R, Zhang S, Zhang S, Wang G. DeepGeO-2V: deep interactive segmentation of multiple organs from head and neck images with lightweight CNNs. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2019;11851:LNCS.619–9.
81. Liew C. The future of radiology augmented with artificial intelligence: a strategy for success. Eur J Radiol. 2018;102:152–6.
82. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. Theranostics. 2019;9(5):1303–22.
83. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Ithaca: Cornell University Library; 2018.
84. Mathur PMF, Burns MLMP. Artificial intelligence in critical care. Int Anesthesiol Clin. 2019;57(2):89–102.
85. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89–94.
86. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies in medical imaging. BMJ. 2020;366:E2671.
87. O’Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Prediction of amyloid pathology in cognitively unimpaired individuals using voxel-wise analysis of longitudinal structural brain MRI. Alzheimer's Res Therapy. 2019;11(1):66.
88. Oikonomou EK, Siddique M, Antoniades C. Artificial intelligence in medical imaging: a radiomic guide to precision phenotyping of cardiovascular disease. Cardiovasc Res. 2020;106.66.
89. Olthof AW, van Ooijen PAA, Rezaadeh Mehrz MH. Promises of artificial intelligence in neuroradiology: a systematic technographic review. Neuroradiology. 2020;62(1):1265–78.
90. Papanikolaou N, Matos C, Koh DM. How to develop a meaningful radiomic signature for clinical use in oncologic patients. Cancer Imaging. 2020;20(1):33.
91. Peek JC, Bernhoffer M, Wiestler B, Goldberg T, Cremers D, Rost B, et al. Radiomics in radiooncology—challenging the medical physicist. Phys Med. 2018;48:27–36.
92. Peek JC, Wiestler B, Combs SE. Image-guided radiooncology: the potential of radiomics in clinical application. Rec Results Cancer Res. 2020;2016:773–94.
93. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur Radiol Exp. 2018;1(1):35.
94. Petrone PM, Casamitjana A, Falcon C, Artigues M, Operto G, Cacciaglia A, et al. Application of machine learning in radiology practice and research. Am J Roentgenol. 2017;208(4):754–60.
95. Paez A. Grey literature: an important resource in systematic reviews. J Evid Based Med. 2017;6:66.
96. Peterson TA, Besser AH, Sall S, Golden D, Hsiao A. Machine learning and deep neural networks in thoracic and cardiovascular imaging. J Thorac Imaging. 2019;34(3):192–201.
97. Sabia L, Biswas M, Uppal SC, Godin J, Hildebrandt M. The present and future of deep learning in radiology: Eur J Radiol. 2019;114:14–24.
98. Shen FX, Wolf SM, Gonzalez RG, Garwood M. Ethical issues posed by field research using highly portable and cloud-enabled neuroimaging. Neuron. 2020;105(5):771–5.
99. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. Eur J Nucl Med Mol Imaging. 2019;46(13):2656–72.
100. Starmans MPA, Voort SR, Tovar AMC, Veenland JF, Klein S, Niessen WJ. Radiomics. In: Handbook of medical image computing and computer assisted intervention. 2019. P. 429–56.
101. Takahashi R, Kajikawa Y. Computer-aided detection: cost effectiveness analysis with learning model. In: 2017 Portland international conference on management of engineering and technology (PICMET). 2017. p. 1–8.
102. Tesche C, De Cecco CN, Albiecht MH, Duguay TM, Bayer RR II, Litwin SE, et al. Coronary CT angiography-derived fractional flow reserve. Radiology. 2017;285(1):17–33.
103. Timmereren J, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—how-to guide and critical reflection. Insights Imaging. 2020;11(1):66.
104. Van Timmereren J, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—how-to guide and critical reflection. Insights Imaging. 2020;11(1):66.
105. Xie Q, Faust K, Van Ommeren R, Sheikh A, Djuric U, Diamandis P. Deep learning for image analysis: personalizing medicine closer to the point of care. Crit Rev Clin Lab Sci. 2019;56(1):61–73.
106. Xie Y, Gunasekaran DV, Balasak S, Keane PA, Sim DA, Bachmann LM, et al. Health economic and safety considerations for artificial intelligence applications in digital retinopathy screening. Transl Vis Sci Technol. 2020;9(2):22.
107. Yuan J, Fan Y, Li X, Chen C, Li D, Hong Y, et al. Research on the practical classification and privacy protection of CT images of parotid tumors based on RestNet50 model. J Phys Conf Ser. 2020;615:1767.
108. Zaharchuk G. Next generation research applications for hybrid PET/MR and PET/CT imaging using deep learning. Eur J Nucl Med Mol Imaging. 2019;46(13):2700–7.
109. Keel S, Lee PY, Scheetz J, Li Z, Kotowicz MA, Maciasri RJ, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Sci Rep. 2018;8(1):4430.
110. Kohli M, Prevedello LM, Filice RW, Geis JR. Implementing machine learning in radiology practice and research. Am J Roentgenol. 2017;208(4):575–60.
111. Lin H, Li R, Liu Z, Chen J, Yang Y, Chen H, et al. Diagnostic efficacy and therapeutic decision-making capacity of an artificial intelligence platform for childhood cataracts in eye clinics: a multicentre randomized controlled trial. eClinicalMedicine. 2019;5:52–9.
112. Pesapane F, Suter MB, Codari M, Patella F, Volonté C, Sardanelli F. Chapter 52—Regulatory issues for artificial intelligence in medicine. In: Faintuch J, Faintuch S, editors. Precision Medicine for Investigators, Practitioners and Providers. New York: Academic Press; 2020. p. 533–43.
113. Pesapane F, Volonté C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. Insights Imaging. 2018;9(5):745–53.
114. Saldan NM, Banja JD, Melzer CC. Ethical considerations in artificial intelligence. Eur J Radiol. 2020;122:108768.
115. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfet J, et al. Canadian Association of radiologists white paper on artificial intelligence in radiology. Can Assoc Radiol J. 2018;69(2):120–35.
116. Grall L, Ieraci L, Hosseinikhah N, Argent-Katwala M. Artificial intelligence in outcomes research: a systematic scoping review. Expert Rev Pharmaacoceon Outcomes Res. 2021;1:61–3.
117. Kidholm K, Ølholm AM, Birk-Olsen M, Cicchetti A, Fure B, Halmesmäki E, et al. Model for ASsessing the value of AI in medicine. In: Handbook of medical image computing and computer assisted intervention. 2019. P. 429–56.
118. Omomu P, Ducarouge A, Tournier A, Harvey H, Kahn CE, Louvet-de Verchère F, et al. To buy or not to buy—evaluating commercial AI solutions in radiology (the ECLAIR guidelines). Eur Radiol. 2021;31(6):3786–96.
119. Paez A. Grey literature: an important resource in systematic reviews. J Evid Based Med. 2017;6:66.
120. Fasterholdt I, Kjølhede T, Naghavi-Behzad M, Schmidt T, Rautalammi Q, Hildebrandt M, et al. Model for ASsessing the value of AI in medical imaging (MAS-AI). International Journal of Technology Assessment in Health Care. 2022;38(1), e74, 1–6. https://doi.org/10.1017/S0266464222000551.