



## OPEN ACCESS

## EDITED BY

Philip M. Meyers,  
Columbia University, United States

## REVIEWED BY

Michela Gabelloni,  
University of Pisa, Italy  
Matt Wagner,  
University of Toronto, Canada

## \*CORRESPONDENCE

Mark Schweitzer  
✉ mark.schweitzer@med.wayne.edu

RECEIVED 22 January 2023

ACCEPTED 27 April 2023

PUBLISHED 15 May 2023

## CITATION

Khosravi P and Schweitzer M (2023) Artificial intelligence in neuroradiology: a scoping review of some ethical challenges.

Front. Radiol. 3:1149461.

doi: 10.3389/fradi.2023.1149461

## COPYRIGHT

© 2023 Khosravi and Schweitzer. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Artificial intelligence in neuroradiology: a scoping review of some ethical challenges

Pegah Khosravi<sup>1</sup> and Mark Schweitzer<sup>2\*</sup>

<sup>1</sup>Department of Biological Sciences, New York City College of Technology, CUNY, New York City, NY, United States, <sup>2</sup>Office of the Vice President for Health Affairs Office of the Vice President, Wayne State University, Detroit, MI, United States

Artificial intelligence (AI) has great potential to increase accuracy and efficiency in many aspects of neuroradiology. It provides substantial opportunities for insights into brain pathophysiology, developing models to determine treatment decisions, and improving current prognostication as well as diagnostic algorithms. Concurrently, the autonomous use of AI models introduces ethical challenges regarding the scope of informed consent, risks associated with data privacy and protection, potential database biases, as well as responsibility and liability that might potentially arise. In this manuscript, we will first provide a brief overview of AI methods used in neuroradiology and segue into key methodological and ethical challenges. Specifically, we discuss the ethical principles affected by AI approaches to human neuroscience and provisions that might be imposed in this domain to ensure that the benefits of AI frameworks remain in alignment with ethics in research and healthcare in the future.

## KEYWORDS

artificial intelligence, data privacy, deep learning, liability, machine learning, neural network, neuroradiology

## Introduction

Artificial intelligence (AI) leverages software to digitally simulate the problem-solving and decision-making competencies of human intelligence, minimize subjective interference, and potentially outperform human vision in determining the solution to specific problems.

Over the past few decades, neuroscience and AI have become to some degree, sublimated with the development of machine learning (ML) models using the brain circuits as the model for the invention of intelligent artifacts (1).

Neuroscience inspired and then ironically, validated the architecture of various AI algorithms (2) such as artificial neural networks (ANNs). This is a subset of ML approaches and is composed of units that are called artificial neurons which are typically organized into input, hidden, and output layers. One of the most successful ANN-based computational models is termed deep neural networks (DNNs) which consist of multiple hidden layers to learn more informative features and ubiquitous fields, by employing intelligence without explicit programming, to solve predictive problems such as segmentation and classification (3, 4).

Traditional, clinical neuroradiologists identify abnormalities of the spine, head and neck, and spinal cord through pattern abnormalities on MRI and CT (and occasionally other modalities). Radiomics and AI leverage imperceptible variations in images. Some have termed this as seeing the unseen. Research on applying AI to neuroradiology, and all imaging, has rapidly grown over the past decade. The number of scholarly publications related to the development and application of AI to the brain has more recently increased astoundingly (5).

The immense increase in publications on the development of AI models shows that AI is rapidly gaining importance in neuroradiologic research and clinical care. The growth in research is related to the combination of more powerful computing resources as well as more advanced measurement techniques (6–8), combined with advances in imaging sequences. Just a few somewhat basic examples of the application of AI models in neuroradiology include image reconstruction (9), improved image quality (10), lesion segmentation (11), specific identification of hemorrhages (12) and other lesions, as well as patterns in psychiatric disease (depression and schizophrenia) and neurologic disorders (such as Huntington's) (13), among many others. The more recent works on neuroimaging AI have concentrated on prognostication and greater personalization of treatments.

For neuroradiology, a deep learning (DL) model receives image series in the input layer, then the extracted features are analyzed in the hidden layer using various mathematical functions. The output layer encodes the desired outcomes or labeled states (e.g., tumor or normal). The goal of training a DL model is to optimize the network weights so that when a new series of sample images are fed to the trained model as inputs, the probabilities measured at the output are heavily skewed to the correct class (14) (Figure 1).

For the short term, these automatic frameworks will likely serve as decision-support tools that will augment the accuracy and efficiency of neuroradiologists. However, the progress in developing these models has not corresponded with progress in implementation in the clinic (16). This may be related to regulatory, reimbursement issues and perhaps most importantly concerns over the adjudication of liability (17). Relatedly, the implementation of AI methods leads to systemic risks with potentially disastrous consequences and societal implications (16).

In this paper, we contribute to the ethical framework of AI in neuroradiology. Below we present several specific ethical risks of AI, as well as air some principles that might guide its development in a more sustainable and transparent way. In our review, we highlight the ethical challenges that are raised by providing input for the AI models and the output of the established frameworks in neuroradiology (Figure 1). In each section, we also present the strategies that might be useful to tackle these challenges in the future.

## The profit of data

The value of any recorded medical observation - whether it is images, physical examination results, or laboratory findings - primarily lies in how they contribute to that patient's care. However, when the data are collected (and anonymized), this data can also be used to generate useful information for potential research. This information also may eventually have commercial potential. Databases can help us understand disease processes or may generate new diagnostic or treatment algorithms. These databases can be used to assess therapies *in silico*.

Because of the value of medical data, especially images, when conglomerated, those who participate in the health care system have an interest in advocating for their use for the most

beneficial purposes. Hence, the controversy regarding who has the right to control and benefit from collected images, patients, or provider organizations.

There has been a paradigm shift over the last decade with respect to the ethos of data ownership (18). While some authors have advocated for researchers' right to access data (19), others have highlighted embedding trustworthiness into data-sharing efforts for all who participate in this process and benefit from it, including patients, providers, healthcare organizations, and industry (20, 21).

Most AI scientists believe in the ethical responsibility of all patients to share their data to improve patient care, now, and in the future. Thus, these researchers believe the data should be widely available and treated as a form of public good to be used for the benefit of future patients (21, 22). Individuals, in various countries and institutions, can and should have the right to opt out of anonymized data use. However, if that becomes widespread, it worsens another AI ethical issue-database biases (*vide infra*).

## Protection of data privacy

The significance of the diligent and ethical use of human data has been highlighted recently to promote a culture of data sharing for the benefit of the greater population, while also protecting the privacy of individuals (18). Therefore, before medical images can be used for the development of research or a commercial AI algorithm, they are required, in most jurisdictions, to obtain approval from the local ethical committee. An institutional review board (IRB) needs to assess the risks and benefits of the study to the patients. In many cases, existing (retrospective) data is used. Because the patients in this type of study do not need to undergo any additional procedures, explicit (written) informed consent is generally waived. With clinical trials, each primary investigator may need to provide approval to share data on their participants. In the case of a prospective study, where study data are gathered prospectively, written informed consent is necessary. After ethical approval, relevant data needs to be accessed, queried, and deidentified for all the health information (PHI) to meet the health insurance portability and accountability act (HIPAA) requirements in the United States or general data protection regulation (GDPR) in Europe, as well as securely stored. Notably, in certain countries, no formal IRB approval is needed for retrospective data use, and in others, patients sign blanket consent when they enter the local medical system. Although this is legal in those environments, to these authors it creates its own ethical concerns but does decrease database biases.

Other issues specifically related to neuroradiology, such as the presence of personal identifiers, which need to be removed both from the digital imaging and communications in medicine (DICOM), metadata, as well as from the images prior to data sharing (23). For example, surface reconstruction of volume acquisitions of the face and brain may allow re-identification by providing detailed images of the patient and generalization of facial recognition techniques (24). To prevent re-identification, defacing or skull-stripping techniques must be used in

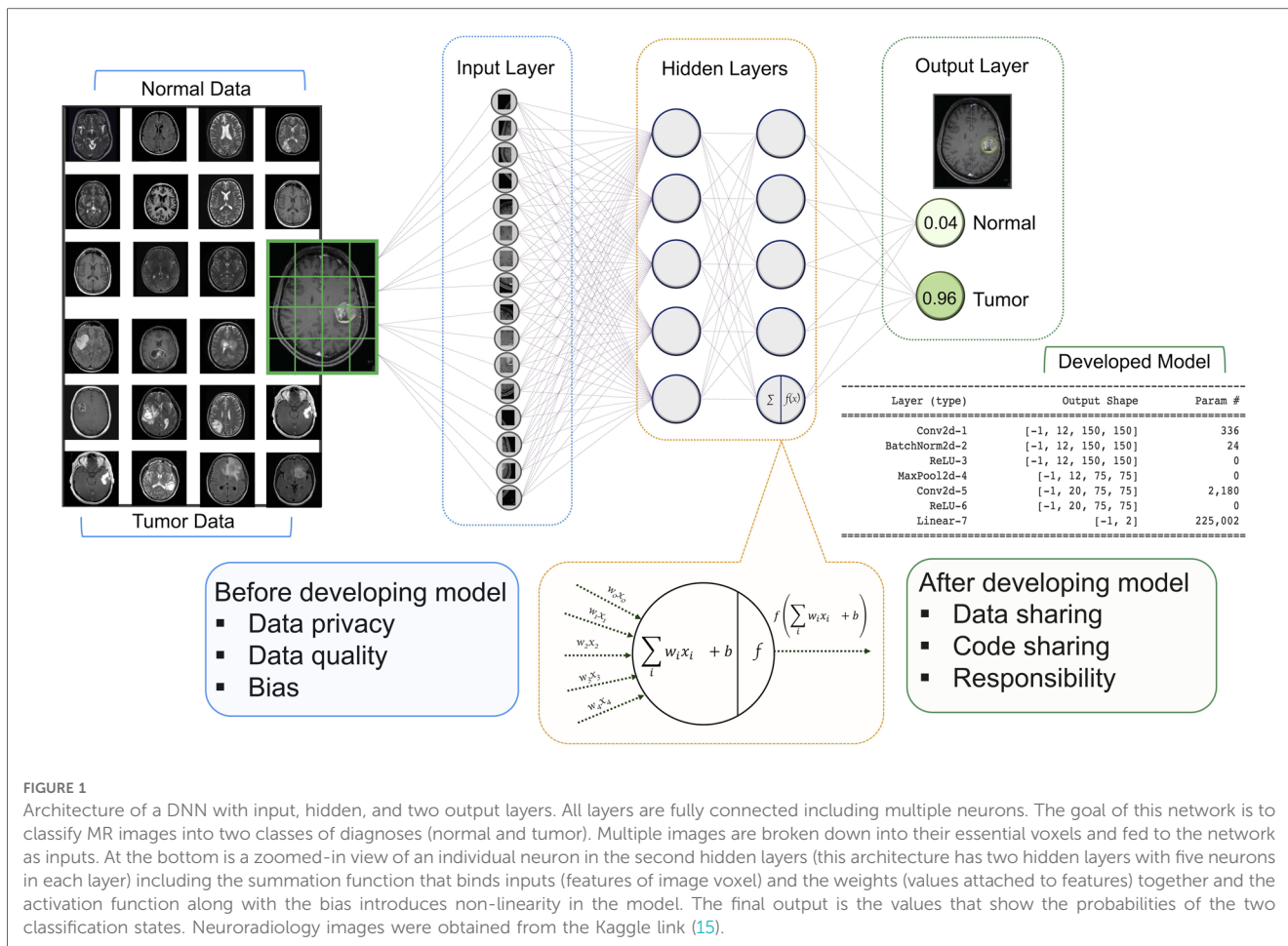


FIGURE 1

Architecture of a DNN with input, hidden, and two output layers. All layers are fully connected including multiple neurons. The goal of this network is to classify MR images into two classes of diagnoses (normal and tumor). Multiple images are broken down into their essential voxels and fed to the network as inputs. At the bottom is a zoomed-in view of an individual neuron in the second hidden layers (this architecture has two hidden layers with five neurons in each layer) including the summation function that binds inputs (features of image voxel) and the weights (values attached to features) together and the activation function along with the bias introduces non-linearity in the model. The final output is the values that show the probabilities of the two classification states. Neuroradiology images were obtained from the Kaggle link (15).

multicenter studies of neuroradiology (25, 26). Also, it is important to remove or blur the surface-based features in the high-resolution MR images to reduce the possibility of re-identification of individuals based on their surface anatomy through pre-processing algorithms. Furthermore, recent studies have shown that the generative adversarial networks (GANs) models and their synthetically generated data can be used to infer the training set membership by an adversary who has access to the entire dataset and some auxiliary information (27). These are somewhat unique issues in neuroimaging research. In this regard, there are software packages available that are able to remove the facial outline from high-resolution radiology images without removing or altering brain tissue (28, 29). Also, a new GAN architecture (privGAN) has been developed to defend membership inference attacks by preventing memorization of the training set to provide protection against this mode of attack (27).

Regarding the increasing usage of digital data in neuroscience, it is important to manage data resources credibly and reliably from the beginning. Data governance is the set of fundamental and policies that illuminate how an organization manages the data assets (30). Data governance is a pivotal module in the analysis of neuroradiology data and includes data stewardship, ownership, policies, and standards (30). The important underlying principle of data governance is the sense that those who manage patient data have an ethical responsibility to manage the data as trustee

for the patients profit, the institution, and the public. Therefore, understanding data governance principles is critical in enabling healthcare professionals to adopt AI technologies effectively and avoid progress of these techniques due to concerns around data security and privacy (31).

Centralized and distributed networks are two organized methods that have been developed for tactical governance of data sharing (32). While in the centralized network the data such as repositories of neuroradiology images from multiple sites are aggregate into single central database, the distributed network model limits the flow of patient data to control the use of information and adhere to applicable legal regulations (32). While the former method improves the consistency of the data, enabling data access is more complex since it needs to adhere to the legal statutes of several organizations. The latter method enables users encounter less barriers related to participant willingness to share data, and legal obstacles. However, they struggle with harmonization of site-specific data before performing analyses at each site to make sure the data are consistent.

### Quality of annotated data

The quality and amount of the annotated images for training AI models are variable, based on the target task. Although using

poor-quality images may lead to poor predictions and assessments, in heterogeneous large quantities, it is known that AI algorithms can be trained on relatively poor-quality images (23, 33). However, knowing the correct label for a given task to correlate with the imaging findings is a critical topic in policy documents and essential for the development of any AI system. In general, imaging data can be labeled in a variety of ways, including image annotations and segmentations. We have an ethical imperative to do AI research well, and one of the requirements is to have an adequate population, to reach conclusions from.

Most currently implemented AI algorithms for medical image classification tasks are based on a supervised learning approach. This means that before an AI algorithm can be trained and tested, the ground truth needs to be defined and attached to the image. The term ground truth typically refers to information acquired from direct observation [either of the images, or the patient (dermatology and surgical exploration), pathological proof, or occasionally clinical follow-up]. For direct observation references standards images are annotated by medical experts—such as neuroradiologists.

Manual labeling is often used in the labeling and annotation of brain imaging data for AI applications. When a relatively small number of images are needed for AI development, medical expert labeling and segmentation may be feasible. However, this approach is time-consuming and costly for large populations, particularly for advanced modalities with numerous images per patient, such as CT, PET, or MRI. Prior to the widespread use of AI, and markedly improved through AI, semiautomated and automated algorithms have become somewhat widely used in labs.

Radiology reports are not created for the development of AI algorithms and the extracted information may contain noise. Although, recently more reports are structured, or protocol-based, more often they are narrative. These narrative reports have previously been assessed through natural language processing programs.

Neural networks can still be relatively robust when trained with noisy labels (34). However, one should be careful when using noisy labels for the development of clinically applicable algorithms because every labeling error could be translated to a decrease in algorithm accuracy. It is estimated that about 20 percent of radiology reports contain noticeable errors due to technical factors or radiologists' specific oversight or misinterpretations (35). Although errors or discrepancies in radiology are inevitable, some can be avoided by appropriate available strategies such as the fusion of radiological and pathological annotation, the use of structured reporting and computer-aided detection, defining quality metrics, and encouraging radiologists to contribute to the collation of these metrics (36).

AI could serve a role to reduce these errors and may help in the more accurate and time-efficient annotation of CT and MRI scans. However, this assumes a sufficiently sized dataset to adequately train the AI models. Furthermore, recent efforts in the automation of the annotation process particularly in neuroimaging data have shown a significant increase in the performance of the annotation process by using AI systems in large scale of data. For example, a study in brain MRI tumor

detection indicated that applying semi-supervised learning to mined image annotations can improve tumor detection performance significantly by achieving an F1 score of 0.95 (37).

There is a trend toward interactive collaboration between AI systems and clinical neuroradiologists. The ability to self-validate and learn from mistakes of AI systems makes the system able to recognize their errors and self-correct their own data sets. On the other hand, external validation by neuroradiologists, manually fine-tuning the AI system, that it has made an error, and allowing the AI system to update its algorithm to avoid errors is helpful for the improvement of these models. Such interactive collaboration annotations have been used effectively for labeling open-source data sets to improve annotation quality while saving radiologists time (38, 39).

## Collaboration of AI and neuroradiologists

Neuroradiology is overall the third most common imaging subspecialty (40), and this subspecialty already heavily utilizes computing and machine technologies. As AI frameworks progress, they can support neuroradiologists by increasing their accuracy and clinical efficiency. However, it is important to recognize that AI applications can also replace some aspects of a neuroradiologist's work and therefore neuroradiologists may have concerns about the progress of AI beyond the role of assistance and replace them. This brings to mind, wider societal concerns about the future role of AI and human work.

A recent review paper explored the application of AI in neuroradiology and identified the majority of these applications provide a supportive role to neuroradiologists rather than replacing their skills. For example, assessing quantitative information such as the volume of hemorrhagic stroke and automatically identifying and highlighting areas of interest such as large vessel occlusions on CT angiogram (41). AI can facilitate and assist neuroradiologists' workflow at several stages. It can extract relevant medical information autonomously and exclude findings not relevant to the investigation. AI can standardize every scan and reduce any artifacts. It can help neuroradiologists by identifying abnormalities that are the most time-sensitive decision and annotating multiple images and other related clinical information at the same time (42).

While these capabilities are very promising, AI systems are not without limitations. For an AI system's algorithm to accurately work, they require large datasets and accurate expert labeling that may not always be possible. Therefore, as Langlotz suggested (43), maybe the neuroradiologists who use AI will replace those who don't keep up to date with AI technologies.

Despite the widespread application of AI, it is important that radiologists remain engaged with AI scientists to both understand the capabilities of existing methods and direct future research in an intelligent way that is supported by the sufficient clinical need to drive widespread adoption (14). An AI-neuroradiologist collaborative model will have high value in potentially improving patient outcomes.

The subject of explicability is crucial as AI models are still considered as black box, due to lack of clarity regarding the data transformations after various passages within the convolutional neural networks (CNNs) despite the mathematical and logical processes (44). Therefore, the close collaboration between neuroradiologists and AI scientists with different expertise aids the development of AI models. While neuroradiologists may annotate images and query the quality of datasets, AI scientists build the algorithms and frameworks. Neuroradiologists who are familiar with digital imaging and informatics have the potential to investigate the black box and accompany the development process, confirming the respect of standards. This collaboration establishes trust and establishes criteria for validating the performance of the AI models. The result of collaboration between AI scientists and neuroradiologists highlights the variety of benefits including the interpretability of the output of the algorithm. Also, while the AI scientists become familiar with the image features from the neuroradiologists' point of view, radiologists are trained to understand how AI works and how to integrate it into practice, how to evaluate its performance as well as recognize the ethical matters involved. In this formal accreditation process, the endpoint is the evaluation of the performance of the AI frameworks, neuroradiologists learning and using the AI systems, and patients trusting the physician using AI tools.

There is insufficient practicality related to the final stages such as reporting and communication compared to the early stages of the AI application (e.g., image pre-processing) (41). This indicates the opportunities for AI scientists and neuroradiologists to develop AI systems in areas beyond image interpretation that can be helpful for communicating the results of medical procedures to patients, particularly in the case that such results could alter the choice of therapy. Patients can benefit from counterfactual recommender systems that learn unbiased policies from data and it can be applied prospectively to support physicians' and patients' management decisions. Using this system have great potential and can be possible in the future when annotated patient data are accessible at the scale (4). Future development of AI applications that integrate software platforms intended for automatic rapid imaging review and provide a communication platform and optimized workflow to multidisciplinary teams will undoubtedly play a key role in more rapid and efficient identification for therapy, resulting in better outcomes in neurological treatments.

## Availability of models and data

The emergence of using AI in neuroradiology research has highlighted the necessity of code sharing because these are key components to facilitating transparent and reproducible experiments in AI research (45). However, despite the recommendations by editorial board members of the radiological society of north america (RSNA) journals (46, 47), recent studies revealed that less than one-third of the articles share code and adequately documented methods (48, 49) and most articles with code sharing by radiographic subspecialty are in neuroradiology (49).

Although these low rates of code sharing and code documentation may be discouraging, it has shown that code sharing over time is an uphill journey (49). Nonetheless, there is room for improvement, which can be facilitated by journals and the peer-review process. For example, reproducible code sharing can be improved by radiology journals through mandatory code and documentation availability upon article submission, reproducibility checks during the peer-review process, and standardized publication of accompanying code repositories and model demos which lead to faster and more collaborative scientific innovation.

In addition to code sharing, data availability is another key component of the reproducibility of AI research studies, because DL models may have variable performance on different neuroimaging datasets (50). However, the rate of data availability is even less than code-sharing (less than one-sixth of studies) (49) due to the challenges related to medical data sharing that we discussed earlier in this review. The latter study showed that the majority of studies that provided data used data from open-source datasets (51) such as TCIA which highlights the importance of these publicly released datasets to research (49). Code and data availability are ethical imperatives for AI research.

## Equal distribution of AI

Medical imaging, including MRI, is one of the most common ways of brain tumor detection. MRI scans such as T2-weighted and post-contrast T1-weighted are preferred since they provide more precise images and provide better visualization of soft tissue, therefore they can be used for brain tumor segmentation (52). CNNs, a class of DNNs, are used as prominent methods for medical image analysis. Therefore, neuroradiologists have become interested in obtaining image features and detecting brain tumors using CNNs to devote less time to screening medical images and more time to image analysis. Using AI tools by neuroradiologists at "the top of their license" is the most ethical construct. Hence, the use of AI to "flag" images, allows highly trained physicians to do what they do best, and what AI might not do best (for now).

Bias is an important ethical theme in research and clinical care. It is quite easy for potential bias to be embedded within algorithms that grow from "selected" data that are used to train algorithms (53, 54). Most research is performed at academic centers, more of this at the most prestigious academic centers. These centers see a biased population that skews towards affluence, education, and often disproportionately small numbers of patients of color. If algorithms are developed on datasets that are under- or over-representative of certain population subgroups, they may exhibit bias when deployed in clinical practice, leading to unequal access to care and potential harm to patients. Using biased data sets not only can potentially cause systemic inequities based on race, gender, and other demographic characteristics, but they may also limit the performance of AI as a diagnostic and treatment tool due to the lack of generalizability (54, 55).

A recently published review by Das et al. (56) investigated the risk of bias in AI data and methods that are used in brain tumor

segmentation. They showed variance in designing the AI architecture and input data types in medical imaging increase the risk of bias. In another study to identify skin cancer from images, the researchers trained a model on a dataset including 129,450 samples, but less than 5% of these images were associated with dark-skinned individuals, thus the performance of the classifier could vary significantly across different populations (57). Furthermore, Larrazabal et al. (58) highlighted the importance of gender balance in medical imaging datasets for training AI systems in computational diagnosis. Using multiple DNN architectures and available image datasets, they showed dramatic changes in performance for underrepresented genders when balance and diversity are not fulfilled.

Some studies suggested using AI itself to potentially mitigate existing bias by reducing human error and biases that are present within healthcare research and databases (59). These studies addressed the issue of bias including building AI systems to reflect current ethical healthcare standards and ensuring a multidisciplinary approach to the deployment of AI (53, 60). To ensure that AI systems are fair and equitable, it is essential to address this challenge by developing algorithms that are trained on diverse datasets, including data from underrepresented populations. Furthermore, it is important to continually monitor AI systems to detect and mitigate any potential biases that may emerge over time. By addressing these challenges, we can help ensure that AI systems in neuroradiology are developed and deployed in a responsible and ethical manner and that they ultimately benefit more patients.

## Liability of the developed AI methods

In this section, we explore a frequent question regarding the application of AI in neuroradiology as to who is responsible for errors that may happen through the process of developing and deploying AI technology. AI algorithms may be susceptible to differences in imaging protocols and variations in patient numbers and characteristics. Thus, there are by the nature of the beast, specific scenarios where these algorithms are less reliable. Therefore, transparent communication about selection criteria and code sharing is required to validate the developed model by external datasets to ensure the generalizability of algorithms in different centers or settings (61).

Despite the success and progress of AI methods, they are ultimately implemented by humans, hence, some consideration of user confidence and trust (62) is important. Also, if AI systems were to fail - as is inevitable - especially if they are involved in medical decision-making, we should be able to determine why and how they failed. Hence, AI processes need to be auditable by authorities, thus enabling legal liability to be assigned to an accountable body (63). The transparency of AI applications in neuroradiology should be considered to ensure that responsibility and accountability remain with human designers or operators (64). To which degree, is liable, becomes an urgent and necessary question.

Several studies recommend that since healthcare professionals are legally and professionally responsible for making decisions in

the patient's health interests, they should be considered responsible for the errors of AI in the healthcare setting, particularly with regard to errors in diagnostic and treatment decisions (65, 66). However, other studies emphasize that it is AI developers' responsibility to ensure the quality of AI technologies, including safety and effectiveness (67, 68). Although a small number of articles suggested commercial strategies for responsible innovation (69), the question is still debatable because AI processes are often complicated to understand and examine the output of AI systems (70).

The use of AI technology in neuroradiology should be cognizant that ultimately it is a person or persons who are responsible for the proper implementation of AI. Part of this responsibility lies in the appropriate implementation and use of guidelines to minimize both medical errors and liability. Guidelines should be implemented to reduce risks and provide reasonable assurance including the well-documented developed methods, research protocols, appropriate large datasets, performance testing, annotation and labeling, user training, and limitations (71). Especially since adopting AI in neuroradiology is increasing and it has enormous potential benefits, implementation of AI in this field requires thoughtful planning and diligent reassessment.

## Case numbers

When we do research, we have a moral obligation to provide the highest quality product we can. That is why IRBs not only look at the risks and benefits but also the research protocol, to ensure that the results will be worthwhile. One of the aspects of this IRB assessment is the size and type of population studied. Similarly in AI research, the database should include a large enough number of subjects to avoid overfitting. In addition, an external validation set is ethically required, to ensure generalizability. The numbers needed in the development set have been a moving target but are appropriately moving towards substantive requirements.

Dataset size is a major driver of bias and is particularly associated with the size of training data; the AI models should be trained on large, heterogenous, annotated data sets (72). Small training brain image datasets lead to overfitting in CNN models which causes diminishing robustness of developed methods due to data-induced bias.

Previous studies also have shown that using a large dataset decreases model bias and yields optimal performance because each MRI technique has different characteristics, therefore, integrating various modalities and techniques yields more accurate results than any single modality. Using multimodality and heterogenous datasets also can handle the overfitting problem regarding the trained model using a specific dataset to make the model generalize for an external validation set (33, 73).

In diseases, or neurologic disorders that are less common, finding large enough datasets may be difficult and may require pooled resources. When providing larger datasets is not possible due to rare diseases and under-represented populations, transfer

learning, and data augmentation can be used to avoid overfitting due to small or limited data sets (74). The studies also presented some recommendations for improving the risk of bias in AI and providing a direction toward selecting appropriate AI attributes. They highlighted the role of some characteristics such as data size, gold standard, DL architecture, evaluation parameters, scientific validation, and clinical evaluation (56) in developing robust AI models. For instance, deep reinforcement learning and deep neuroevolution models have generalized well based on sparse data and successfully used for the evaluation of treatment response in brain metastasis and classification of brain tumors using MR images (75, 76). Nonetheless, for the time being, some assessments of unusual conditions may not be ethically researched and evaluated by AI, due to the limited availability of copious patient data.

## Context

Several ethical themes of the applications of AI in neuroradiology were presented in this review including privacy and quality of data that are used for training AI systems as well as the availability and liability of the developed AI models. In this scoping review, we addressed the ethics of AI within neuroradiology and reviewed overarching ethical concerns about privacy, profit, liability, and bias, each of which is interdependent and mutually reinforcing. Liability, for instance, is a noted concern when considering who is responsible for protecting patient privacy within data-sharing partnerships and for AI errors in patient diagnoses. We note that liability is related to specific laws and legislation, which by definition vary from one to another jurisdiction.

These broad ethical themes of privacy and security, liability, and bias have also been reported in other reviews on the application of AI in healthcare and radiology in general, and neuroradiology in particular. For example, in a review by Murphy et al. (69), the authors discussed ethical issues including responsibility surrounding AI in the field of health and pointed to a critical need for further research into the ethical implications of AI within both global and public health.

In another study, the authors discussed ethical principles including accountability, validity, the risk of neuro-discrimination, and neuro-privacy that are affected by AI approaches to human neuroscience (6). These latter two terms likely will be increasingly part of our conversations going forward.

Another article with a focus on patient data and ownership covered key ethical challenges with recommendations towards a sustainable AI framework that can ensure the application of AI for

radiology is molded into a benevolent rather than malevolent technology (17). The other recent studies highlighted the intersection of data sharing, privacy, and data ownership with specific examples regarding neuroimaging (18). It is therefore clear from all review articles that the ethical challenges in AI should be considered in relation to all people who participate in developing AI technology including neuroradiologists and AI scientists.

In conclusion, the ethical challenges surrounding the application of AI in neuroradiology are complex, and the value of AI in neuroradiology increases by interdisciplinary consideration of the societal and scientific ethics in which AI is being developed to promote more reliable outcomes and allow everyone equal access to the benefits of these promising technology. Issues of privacy, profit, bias, and liability have dominated the ethical discourse to date with regard to AI and health, and there will undoubtedly be more that arise. AI is being developed and implemented worldwide, and thus, a greater concentration of ethical research into AI is required for all applications amidst the tremendous potential that AI carries, it is important to ensure its development and implementation are ethical for everyone.

## Author contributions

The paper was designed and written by PK and MS both co-authors provided feedback and were involved in manuscript revision. All authors contributed to the article and approved the submitted version.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ullman S. Using neuroscience to develop artificial intelligence. *Science*. (2019) 363(6248):692–3. doi: 10.1126/science.aau6595
- Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-Inspired artificial intelligence. *Neuron*. (2017) 95(2):245–58. doi: 10.1016/j.neuron.2017.06.011
- Greenspan H, van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans Med Imaging*. (2016):1153–9. doi: 10.1109/tmi.2016.2553401
- Boehm KM, Khosravi P, Vanguri R, Gao J, Shah SP. Harnessing multimodal data integration to advance precision oncology. *Nat Rev Cancer*. (2022) 22(2):114–26. doi: 10.1038/s41568-021-00408-3

5. Lui Yw, Chang PD, Zaharchuk G, Barboriak DP, Flanders AE, Wintermark M, et al. Artificial intelligence in neuroradiology: current status and future directions. *AJNR Am J Neuroradiol.* (2020) 41(8):E52–9. doi: 10.3174/ajnr.A6681
6. Ienca M, Ignatiadis K. Artificial intelligence in clinical neuroscience: methodological and ethical challenges. *AJOB Neurosci.* (2020) 11(2):77–87. doi: 10.1080/21507740.2020.1740352
7. Richards B, Tsao D, Zador A. The application of artificial intelligence to biology and neuroscience. *Cell.* (2022) 185(15):2640–3. doi: 10.1016/j.cell.2022.06.047
8. Gopinath N. Artificial intelligence and neuroscience: an update on fascinating relationships. *Process Biochem.* (2023) 125:113–20. doi: 10.1016/j.procbio.2022.12.011
9. Zhu B, Liu JZ, Cauley SF, Rosen B, RRosen MS. Image reconstruction by domain-transform manifold learning. *Nature.* (2018) 555(7697):487–92. doi: 10.1038/nature25988
10. Gao Y, Li H, Dong J, Feng G. *A deep convolutional network for medical image super-resolution. 2017 Chinese automation congress (CAC)* [Preprint]. (2017). doi: <https://doi.org/10.1109/cac.2017.8243724>
11. Mehta R, Majumdar A, Sivaswamy J. Brainsegnet: a convolutional neural network architecture for automated segmentation of human brain structures. *J Med Imaging (Bellingham).* (2017) 4(2):024003. doi: 10.1117/1.JMI.4.2.024003
12. Phong TD, Duong HN, Nguyen HT, Trong NT, Nguyen VH, Van Hoa T, et al. *Brain hemorrhage diagnosis by using deep learning. Proceedings of the 2017 international conference on machine learning and soft computing* [Preprint] (2017). doi: <https://doi.org/10.1145/3036290.3036326>
13. Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, et al. Deep learning for neuroimaging: a validation study. *Front Neurosci.* (2014) 8:229. doi: 10.3389/fnins.2014.00229
14. Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep learning in neuroradiology. *Am J Neuroradiol.* (2018) 39(10):1776–84. doi: 10.3174/ajnr.a5543
15. Chakraborty N. Brain MRI Images for Brain Tumor Detection. (2019). Available at: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection> (Accessed January 11, 2023).
16. Goisauf M, Cano Abadía M. Ethics of AI in radiology: a review of ethical and societal implications. *Front Big Data.* (2022) 5:850383. doi: 10.3389/fdata.2022.850383
17. Mudgal KS, Das N. The ethical adoption of artificial intelligence in radiology. *BJR Open.* (2020) 2(1):20190020. doi: 10.1259/bjro.20190020
18. White T, Blok E, Calhoun VD. Data sharing and privacy issues in neuroimaging research: opportunities, obstacles, challenges, and monsters under the bed. *Hum Brain Mapp.* (2022) 43(1):278–91. doi: 10.1002/hbm.25120
19. Harris TL, Wyndham JM. Data rights and responsibilities: a human rights perspective on data sharing. *J Empir Res Hum Res Ethics.* (2015) 10(3):334–7. doi: 10.1177/1556264615591558
20. McGuire AL, Majumder MA, Villanueva AG, Bardill J, Bollinger JM, Boerwinkle E, et al. Importance of participant-centricity and trust for a sustainable medical information commons. *J Law, Med Ethics.* (2019) 47(1):12–20. doi: 10.1177/107311051984048
21. Larson DB, Magnus DC, Lungren MP, Shah N, HLanglotz CP. Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. *Radiology.* (2020) 295(3):675–82. doi: 10.1148/radiol.2020192536
22. Krupinski EA. An ethics framework for clinical imaging data sharing and the greater good. *Radiology.* (2020) 295(3):683–4. doi: 10.1148/radiol.2020200416
23. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. *Radiology.* (2020) 295(1):4–15. doi: 10.1148/radiol.2020192224
24. SFR-IA Group, CERF and French Radiology Community. Artificial intelligence and medical imaging 2018: french radiology community white paper. *Diagn Interv Imaging.* (2018) 99(11):727–42. doi: 10.1016/j.diii.2018.10.003
25. Chen JJ-S, Juluru K, Morgan T, Moffitt R, Siddiqui K, MSiegel EL. Implications of surface-rendered facial CT images in patient privacy. *AJR Am J Roentgenol.* (2014) 202(6):1267–71. doi: 10.2214/AJR.13.10608
26. Kalavathi P, Surya Prasath VB. Methods on skull stripping of MRI head scan images—a review. *J Digit Imaging.* (2016) 29:365–79. doi: 10.1007/s10278-015-9847-8
27. Mukherjee S, Xu Y, Trivedi A, Patowary N, Ferres JL. privGAN: protecting GANs from membership inference attacks at low cost to utility. *Proc Priv Enhancing Technol.* (2021) 2021(3):142–63. doi: 10.2478/popets-2021-0041
28. Milchenko M, Marcus D. Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics.* (2013) 11:65–75. doi: 10.1007/s12021-012-9160-3
29. Theyers AE, Zamyadi M, O'Reilly M, Bartha R, Symons S, MacQueen GM, et al. Multisite comparison of MRI defacing software across multiple cohorts. *Front Psychiatry.* (2021) 12:617997. doi: 10.3389/fpsy.2021.617997
30. Monah SR, Wagner MW, Biswas A, Khalvati F, Erdman LE, Amirabadi A, et al. Data governance functions to support responsible data stewardship in pediatric radiology research studies using artificial intelligence. *Pediatr Radiol.* (2022) 52(11):2111–9. doi: 10.1007/s00247-022-05427-2
31. Wiljer D, Hakim Z. Developing an artificial intelligence-enabled health care practice: rewiring health care professions for better care. *J Med Imaging Radiat Sci.* (2019) 50(4 Suppl 2):S8–S14. doi: 10.1016/j.jmir.2019.09.010
32. McGraw D, Leiter A. Pathways to success for multi-site clinical data research. *eGEMS.* (2013) 9:13. doi: 10.13063/2327-9214.1041
33. Khosravi P, Lysandrou M, Eljalby M, Li Q, Kazemi E, Zisimopoulos P, et al. A deep learning approach to diagnostic classification of prostate cancer using pathology–radiology fusion. *J Magn Reson Imaging.* (2021) 54(2):462–71. doi: 10.1002/jmri.27599
34. Agarwal V, Podchiyska T, Banda JM, Goel V, Leung TI, Minty EP, et al. Learning statistical models of phenotypes using noisy labeled training data. *J Am Med Inform Assoc.* (2016) 23(6):1166–73. doi: 10.1093/jamia/ocw028
35. Brady A, Laoide R.Ó, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. *Ulster Med J.* (2012) 81(1):3–9. PMID: PMC3609674; PMID: 23536732
36. Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging.* (2017) 9:171–82. doi: 10.1007/s13244-016-0534-1
37. Swinburne NC, Yadav V, Kim J, Choi YR, Gutman DC, Yang JT, et al. Semisupervised training of a brain MRI tumor detection model using mined annotations. *Radiology.* (2022) 303(1):80–9. doi: 10.1148/radiol.210817
38. Folio LR, Machado LB, Dwyer AJ. Multimedia-enhanced radiology reports: concept, components, and challenges. *Radiographics.* (2018) 38(2):462–82. doi: 10.1148/rg.2017170047
39. Yan K, Wang X, Lu L, Summers RM. Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *J Med Imaging.* (2018) 5(3):036501. doi: 10.1117/1.JMI.5.3.036501
40. Bhuva S, Parvizi N. Factors affecting radiology subspecialty choice among radiology registrars in the UK: a national survey. *Clin Radiol.* (2016) 71:S5. doi: 10.1016/j.crad.2016.06.038
41. Olthof AW, van Ooijen PMA, Rezade Mehrizi MH. Promises of artificial intelligence in neuroradiology: a systematic technographic review. *Neuroradiology.* (2020) 62(10):1265–78. doi: 10.1007/s00234-020-02424-w
42. Cohen RY, Sodickson AD. An orchestration platform that puts radiologists in the driver's seat of AI innovation: a methodological approach. *J Digit Imaging.* (2022) 36:1–15. doi: 10.1007/s10278-022-00649-0
43. Langlotz CP. Will artificial intelligence replace radiologists? *Radiol Artif Intell.* (2019) 1(3):e190058. doi: 10.1148/ryai.2019190058
44. Neri E, Coppola F, Miele V, Bibbolino C, Grassi R. Artificial intelligence: who is responsible for the diagnosis? *Radiol Med.* (2020) 125:517–21. doi: 10.1007/s11547-020-01135-9
45. Kitamura FC, Pan I, Kline TL. Reproducible artificial intelligence research requires open communication of complete source code. *Radiol Artif Intell.* (2020) 2(4):e200060. doi: 10.1148/ryai.2020200060
46. Bluemke DA, Moy L, Bredella MA, Ertl-Wagner BB, Fowler KJ, Goh VJ, et al. Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers—from the *Radiology* editorial board. *Radiology.* (2020) 294(3):487–9. doi: 10.1148/radiol.2019192515
47. Mongan J, Moy L, Kahn CE Jr. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell.* (2020) 2(2):e200029. doi: 10.1148/ryai.2020200029
48. Gundersen OE, Kjensmo S. *State of the art: reproducibility in artificial intelligence. Proceedings of the AAAI conference on artificial intelligence.* (2018). doi: <https://doi.org/10.1609/aaai.v32i1.11503>
49. Venkatesh K, Santomartino SM, Sulam J, Yi PH. Code and data sharing practices in the radiology artificial intelligence literature: a meta-research study. *Radiol Artif Intell.* (2022) 4(5):e220081. doi: 10.1148/ryai.220081
50. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Med.* (2018) 15(11):e1002683. doi: 10.1371/journal.pmed.1002683
51. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging.* (2013) 26(6):1045–57. doi: 10.1007/s10278-013-9622-7
52. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging.* (2015) 34(10):1993–2024. doi: 10.1109/TMI.2014.2377694
53. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med.* (2018) 378(11):981–3. doi: 10.1056/NEJMp1714229
54. Ching T, Himmelstein DS, Baulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface.* (2018) 15(141):20170387. doi: 10.1098/rsif.2017.0387
55. Senders JT, Zaki MM, Karhade AV, Chang B, Gormley WB, Broekman ML, et al. An introduction and overview of machine learning in neurosurgical care. *Acta Neurochir.* (2018) 160(1):29–38. doi: 10.1007/s00701-017-3385-8



56. Das S, Nayak GK, Saba L, Kalra M, Suri JS, Saxena S. An artificial intelligence framework and its bias for brain tumor segmentation: a narrative review. *Comput Biol Med.* (2022) 143:105273. doi: 10.1016/j.combiomed.2022.105273
57. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* (2017) 542(7639):115–8. doi: 10.1038/nature21056
58. Larrazabal AJ, Nieto N, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proc Natl Acad Sci U S A.* (2020) 117(23):12592–4. doi: 10.1073/pnas.1919012117
59. Markowitz A, Błaszczewicz K, Montag C, Switala C, Schlaepfer TE. Psychoinformatics: big data shaping modern psychometrics. *Med Hypotheses.* (2014) 82(4):405–11. doi: 10.1016/j.mehy.2013.11.030
60. Howard A, Borenstein J. The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. *Sci Eng Ethics.* (2018) 24(5):1521–36. doi: 10.1007/s11948-017-9975-2
61. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* (2019) 25(1):44–56. doi: 10.1038/s41591-018-0300-7
62. Ghoshal B, Tucker A, Sanghera B, Wong WL. *Estimating uncertainty in deep learning for reporting confidence to clinicians when segmenting nuclei image data.* 2019 IEEE 32nd international symposium on computer-based medical systems (CBMS) [Preprint]. (2019). doi: 10.1109/cbms.2019.00072
63. Akinci D'Antonoli T. Ethical considerations for artificial intelligence: an overview of the current radiology landscape. *Diagn Interv Radiol.* (2020) 26(5):504–11. doi: 10.5152/dir.2020.19279
64. Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Can Assoc Radiol J.* (2019) 70(4):329–34. doi: 10.1016/j.carj.2019.08.010
65. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med.* (2014) 62(1):1–10. doi: 10.1016/j.artmed.2014.06.004
66. Balthazar P, Harri P, Prater A, Safdar NM. Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. *J Am Coll Radiol.* (2018) 15(3 Pt B):580–6. doi: 10.1016/j.jacr.2017.11.035
67. Yuste R, Goering S, Arcas BAY, Bi G, Carmena JM, Carter A, et al. Four ethical priorities for neurotechnologies and AI. *Nature.* (2017) 551(7679):159–63. doi: 10.1038/551159a
68. O'Brolcháin F. Robots and people with dementia: unintended consequences and moral hazard. *Nurs Ethics.* (2019) 26(4):962–72. doi: 10.1177/0969733017742960
69. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics.* (2021) 22(1):1–17. doi: 10.1186/s12910-021-00577-8
70. Mesko B. The role of artificial intelligence in precision medicine. *Exp Rev Precision Med Drug Dev.* (2017) 2(5):239–41. doi: 10.1080/23808993.2017.1380516
71. Jaremko JL, Azar M, Bromwich R, Lum A, Alicia Cheong LH, Gibert M, et al. Canadian association of radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *Can Assoc Radiol J.* (2019) 70(2):107–18. doi: 10.1016/j.carj.2019.03.001
72. Zou J, Schiebinger L. AI Can be sexist and racist — it's time to make it fair. *Nature.* (2018) 559:324–6. doi: 10.1038/d41586-018-05707-8
73. Eche T, Schwartz LH, Mokrane F-Z, Derclé L. Toward generalizability in the deployment of artificial intelligence in radiology: role of computation stress testing to overcome underspecification. *Radiol Artif Intell.* (2021) 3(6):e210097. doi: 10.1148/ryai.2021210097
74. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine.* (2018) 27:317–28. doi: 10.1016/j.ebiom.2017.12.026
75. Stember J, Shalu H. Deep reinforcement learning classification of brain tumors on MRI. *Innov Med Healthc.* (2022) 308:119–28. doi: 10.1007/978-981-19-3440-7\_11
76. Stember JN, Young RJ, Shalu H. Direct evaluation of treatment response in brain metastatic disease with deep neuroevolution. *J Digit Imaging.* (2022) 36:1–11. doi: 10.1007/s10278-022-00725-5