

Received:
23 March 2023

Revised:
13 September 2023

Accepted:
27 September 2023

Published online:
18 October 2023

Cite this article as:

Landry G, Kurz C, Traverso A. The role of artificial intelligence in radiotherapy clinical practice. *BJR Open* (2023) 10.1259/bjro.20230030.

REVIEW ARTICLE

The role of artificial intelligence in radiotherapy clinical practice

^{1,2,3}GUILLAUME LANDRY, ¹CHRISTOPHER KURZ and ^{4,5}ALBERTO TRAVERSO

¹Department of Radiation Oncology, LMU University Hospital, LMU Munich, Munich, Germany

²German Cancer Consortium (DKTK), partner site Munich, a partnership between DKFZ and LMU University Hospital, Munich, Germany

³Bavarian Cancer Research Center (BZKF), Munich, Germany

⁴Department of Radiation Oncology (Maastricht Clinic), School for Oncology and Reproduction (GROW), Maastricht University Medical Center, Maastricht, Netherlands

⁵School of Medicine, Vita-Salute San Raffaele University, Milan, Italy

Address correspondence to: Prof Guillaume Landry
E-mail: guillaume.landry@med.uni-muenchen.de

ABSTRACT

This review article visits the current state of artificial intelligence (AI) in radiotherapy clinical practice. We will discuss how AI has a place in the modern radiotherapy workflow at the level of automatic segmentation and planning, two applications which have seen real-work implementation. A special emphasis will be placed on the role AI can play in online adaptive radiotherapy, such as performed at MR-linacs, where online plan adaptation is a procedure which could benefit from automation to reduce on-couch time for patients. Pseudo-CT generation and AI for motion tracking will be introduced in the scope of online adaptive radiotherapy as well. We further discuss the use of AI for decision-making and response assessment, for example for personalized prescription and treatment selection, risk stratification for outcomes and toxicities, and AI for quantitative imaging and response assessment. Finally, the challenges of generalizability and ethical aspects will be covered. With this, we provide a comprehensive overview of the current and future applications of AI in radiotherapy.

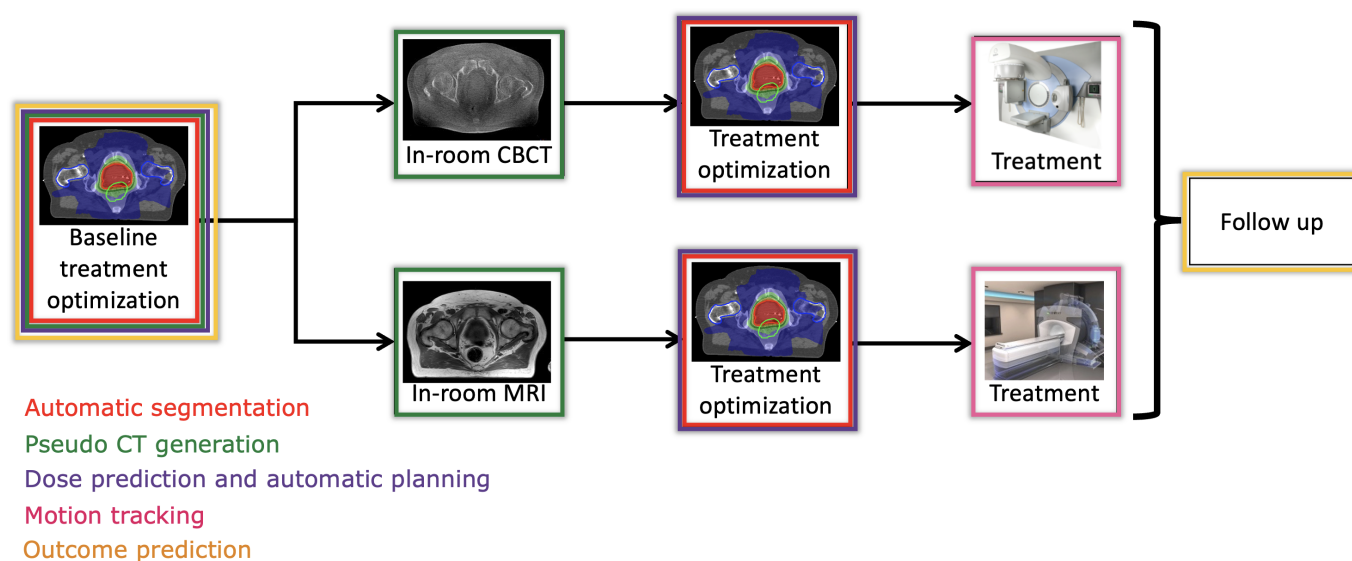
BACKGROUND

Machine-learning techniques, widely referred to as artificial intelligence (AI), are supporting automation in several industries including healthcare, and in particular for this review article, in radiation therapy (RT). Techniques such as deep learning (DL), based on convolutional neural networks (CNNs), have found great success in some applications. This is aided by the relatively large volume of annotated imaging data generated in clinical practice in RT, for example organ segmentations on computed tomography (CT) or magnetic resonance (MR) images. While much remains at the research level, some applications, such as automatic segmentation and automatic planning, have begun seeing clinical use. Additionally, new developments in MR-guided RT (MRgRT) at MR-linacs have increased the frequency of online plan adaptation, which brings a greater need for automation and speed, as well as some specific tasks such as converting MR images into pseudo-CT images. Some machine-learning applications in RT have even preceded the recent wave of enthusiasm for AI, such as decision-making and response assessment using AI applied to medical imaging (radiomics), and have contributed to a culture of collecting datasets in RT. Such data collection

brings related challenges of generalizability of AI methods, privacy and ethical concerns. For automated medical image analysis, both machine learning (e.g., radiomics) and deep learning (e.g., deep radiomics) will be impacted by the interobserver variability in the definition of the Regions of Interest (ROIs) from which features are extracted from. Nevertheless, the published literature strongly supports the use of semi-automated methods for the contouring of ROIs. Recent applications in deep learning have bypassed the problem by avoiding ROIs delineations and using bounding boxes around a selected part of the image, or even using as input the whole 3D image. However, this approach may open issues such as: risk of overfitting, introduction of non-relevant parts of the image as input (i.e., noise [background] vs signal), and higher computational costs. We strongly suggest, when possible, to benchmark the stability of the results against interobserver variability by collecting data with multiple operators or using publicly available datasets.

This review article will cover the current state of AI in clinical practice and provide an overview of potential future applications and challenges. Figure 1 summarizes where AI may play a role in the RT workflow, whether based on

Figure 1. Overview of the modern RT workflow. The top arm represents cone beam computed tomography (CBCT)-guided RT, while the bottom arm represents MRgRT. Various areas of application of AI are highlighted to indicate in which part of the workflow they come into play: automatic segmentation, pseudo-CT generation, dose prediction and automatic planning, motion tracking and outcome prediction.



cone beam computed tomography (CBCT) image guidance or MRgRT.

AI IN CURRENT RADIOTHERAPY WORKFLOWS

The most common application of AI in RT is volume of interest segmentation. For this task, the U-shaped convolutional neural network (U-net) has become the workhorse and consistently produces clinically acceptable results. The method has evolved from 2D segmentation to 3D and has seen the development of dedicated loss functions for model training, such as the Dice similarity coefficient (DSC) or focal loss.¹ By now, state-of-the-art models come with self-configuration pipelines which determine which hyperparameters are optimal for a segmentation task.² AI auto-segmentation is a reality in several clinics, and multiple vendors now provide commercial solutions for CT images (see for example references^{3,4}). Several clinical trials have been initiated to explore auto-segmentation algorithms, *e.g.*, for brain (NCT05093751) or lung lesion segmentation (NCT04164186, NCT05775068). Advantages of automatic segmentation include the reduction of variability among observers⁵ and more importantly time savings in clinical practice. Currently, it is accepted that automatic segmentation is not to be used in an unsupervised manner, but that performing correction of automatic contours saves time compared to a solely manual workflow.^{6,7} Evaluation of the benefits of AI segmentation make use of a set of metrics which include classical geometrical analysis such as a DSC and Hausdorff distance, but also expert evaluation using a grading scale.⁸ The geometric metrics do not readily reflect the main clinical advantage, namely, time saving. For this, besides directly recording correction times, metrics such as surface DSC and added path length (APL) may be useful, and Vaassen *et al*⁷ report that APL correlates the most with time savings.

Another area where AI has made inroads into clinical practice is at the level of automated treatment planning. This part of the

workflow will dramatically reduce the time required for an RT technician to deliver a clinically acceptable treatment plan and is thus an area where efforts toward AI automation have been made.^{9,10} One approach based on machine-learning (ML, referring here to models such as support vector machines or random forests) has been knowledge-based planning,¹¹ where features such as organ and target distances or number of beams are passed to a ML model predicting characteristics of the dose distribution, such as dose volume histograms (DVH) points. These can then be used in an optimizer as objectives, leading to an automatic procedure. More recently, DL methods, again using U-nets, have been used to predict an entire dose distribution instead of a limited set of DVH points, which allows to steer the optimizer to directly achieve that dose distribution.¹² Interestingly, in a study where knowledge-based planning was used clinically, the authors reported that 61% of automatic treatment plans were selected. This was surprising since in a simulated environment, a higher percentage of automatic plans (83%) had been selected, indicating that there remain hurdles to the adoption of this technology in clinical routine.

AI FOR ONLINE ADAPTIVE RADIOTHERAPY

Online adaptive radiotherapy has been made feasible thanks to the clinical introduction of MR-linacs such as the 0.35 T ViewRay MRidian¹³ or 1.5 T Elekta Unity,¹⁴ where MR images are used to perform online plan adaptation, typically in a hypo-fractionated setting.^{15,16} In parallel, CBCT-based systems making use of AI, such as the Varian Ethos,¹⁷ have also recently allowed for clinical online adaptive radiotherapy. Whether CBCT- or MR-based, the online adaptive workflow is one of the areas of radiotherapy with the highest demand for automation. With the patient on the treatment couch, it is highly desirable to accelerate the various steps required to perform plan adaptation, namely, pseudo-CT generation, delineation and plan optimization. At modern MR-linacs, relatively long fraction durations are reported,

ranging from 30 min up to 70 min.^{18–20} Delineation can take up to 30 min and currently the main bottle-neck of the online adaptation process.²¹

Several studies have investigated automatic segmentation specific to MRgRT,^{5,22–29} some of them applied clinically.³⁰ While most segmentation algorithms employ morphological MRI sequences, also studies making use of diffusion-weighted imaging exist.³¹ This is well summarized in the review by Cusumano et al.³² Given the richness of data available, the online adaptive workflow also brings opportunities for novel approaches to automatic segmentation. For example, it may be possible to exploit patient-specific training approaches,³³ where we assume that high-quality, physician-approved planning contours may be available at the time of plan adaptation for a given fraction. In that case, a conventional baseline model trained on a population dataset could be refined on the planning data of a single patient.⁸ Alternatively, the training could be refined as a function of added fraction images.³⁴ Others have trained models from scratch on planning images of an individual patient only,^{35,36} however issues of robustness may arise in this case, especially in the presence of pronounced anatomical changes. In general, when training or fine-tuning models in a patient-specific fashion, care has to be taken to ensure that the model is not overly matched to the planning image, to the detriment of performance on fraction images. One approach is to select a set of patient datasets where fraction images are used as a validation set to monitor the patient-specific training. In this way, it can be ensured that the model performance is optimal on fraction images by selecting an optimal augmentation strategy and training epoch, which will be subsequently fixed for all new patients.⁸ Another issue here is that delineation errors on the planning image may be propagated to the fraction images; however, the availability of more time for planning segmentation should mitigate this. Another way to exploit planning segmentations is to use AI-based DIR, as done in Eppenhof et al for CTV segmentation.³⁷ While they achieved convincing results for CTV, limited evidence is available as to the feasibility of this approach for larger deformations for organs such as the bladder or rectum. The studies listed above were focused on MRgRT; however, similar approaches should be applicable to CBCT-based online adaptive RT. In the latter case, however, additional challenges might arise due to the considerably lower soft tissue contrast of CBCT images and the lack of available ground-truth segmentation on clinical CBCT images. In consequence, CBCT segmentation might be combined with domain adaptation to support auto-segmentation.³⁸

For pseudo-CT generation, also representing a domain adaptation task, several well-established deep-learning methods have by now been identified either for CBCT to CT or MRI to CT conversion, as well reviewed by Spadea et al³⁹ or Gurney-Champion et al for MRI.⁴⁰ The most common approach is to use a U-net, as used by Han et al for MRI to CT⁴¹ and Kida et al for CBCT to CT.⁴² For this approach, paired data are required. This functions well for relatively static anatomies such as the head and neck or brain, where deformable image registration (DIR) may allow correcting slight anatomical mismatches between CBCT/MR and CT scans.⁴² Alternatively, unpaired training may

be used for anatomical sites where DIR may not be sufficient to correct for anatomical variations. For this, generative adversarial networks (GAN) provide an elegant solution.⁴³ The cycleGAN architecture in particular allows to use unpaired data, affording flexibility when collecting datasets for network training. Kurz et al applied this to CBCT to CT conversion with success.⁴⁴ A potential drawback of these generative models is that they might suffer from reduced geometric fidelity, such that additional cross-modality loss terms constraining the networks might be required.^{45,46}

Another area where AI may play a role in online adaptive MRgRT is for motion tracking on cine-MRI. At the 0.35 T MRIdian MR-linac for example, 2D cine-MRI images are acquired at each fraction for beam gating at a frequency of either 4 Hz or 8 Hz. While gating is a very robust way of ensuring target coverage, it does come at the cost of increased treatment times, with duty cycles as low as 20%.⁴⁷ An alternative which has seen recent interest for MRgRT is multi-leaf collimator (MLC) tumor tracking,⁴⁸ which can be achieved by rigidly or deformably shifting the MLC to match the tumor centroid displacement and shape as observed on 2D cine-MRI. Recent proof-of-principle work at the 1.5 T Unity MR-linac showed that MLC tracking should be feasible,^{49,50} and the Australian MR-linac project has even demonstrated tracking multiple targets.⁵¹ Optimal MLC tracking requires target tracking⁵² and accounting for system latency, which has been reported to be about 350 ms for different MR-linacs.^{51,53,54} One approach to account for latency is motion prediction, and AI methods such as long-short-term memory (LSTM) networks have been shown to outperform non-AI methods for this.^{55,56} Another interesting approach for motion tracking itself would be again to use DIR networks to deform a key frame to a moving frame.⁵⁷

DECISION-MAKING AND RESPONSE ASSESSMENT

Advances in RT delivery techniques and the advent of promising new therapies such as immunotherapy or targeted therapies have created a sort of “explosion of treatments”,⁵⁸ from which doctors and patients have to choose. Not only the portfolio of therapies has increased, but many of these treatments interact among each other (*e.g.*, RT enhancing immune response). Personalized medicine refers to choosing the optimal treatment strategy and regimen based on each single patient tumor fingerprints, which are extracted using data mining techniques based on routinely acquired data related to the medical history of the patients, such as electronic health records, genetic testing, and medical imaging.⁵⁹ AI offers a diverse choice of multiple data analytics algorithms that can support personalized medicine.

PERSONALIZED PRESCRIPTION AND TREATMENT SELECTION

Many treatment indications in RT (and oncology in general) strongly rely on the results obtained from randomized controlled clinical trials (RCT). If on one hand, RCTs are recognized for their capability of isolating confounding factors during the comparison of treatments (‘control groups’); on the other hand, the “one size does not fit all” is a well-known limitation of RCTs.

The upcoming challenge in personalized medicine is to use the knowledge derived from RCTs but adapted on a patient-to-patient basis using routinely collected data during the patient journey: real-world evidence.⁶⁰ AI algorithms can support this practice, given their capability of mining a vast amount of unstructured data and find correlations with an outcome of interest. It is important to point out that the role of AI in personalized prescription and treatment selection is not to replace RCTs, but rather suggesting, for example, one among two treatments that have shown comparable outcomes (*e.g.*, local control, survival) but will provide for example less toxicities. These AI algorithms are often referred to as decision support systems (DSS).⁶¹ The most optimal DSS will be based on an estimate of the probability of the outcome (*e.g.*, treatment response) *vs* the risk of unwanted toxicities (*e.g.*, RT side-effects). However, it is difficult to use AI to change the dose prescribed to a tumor, given the fact that dose prescriptions are determined by strict clinical guidelines and considering also that the relation among changes in the dose released to a biological tissue and the resulting DNA changes are still under investigation. At the time of writing treatment decisions and clinical constraints on doses to be prescribed are based on the NTCP (Normal Tissue Control Probability) and TCP (Tumor Control Probability) models. However, the research community is investigating how AI can complement and enhance the above models by introducing tumor- and host(organ at risk)-derived fingerprints. An advanced example is the study by Scott et al, where the investigators validated a genomic-adjusted radiation dose, which is a method that accounts for biological heterogeneity and can be used to predict optimal RT dose for an individual patient.⁶² They have shown that using data from the RTOG 0617 trial, 80% of the patients could have been exposed to unnecessary dose escalation without benefit to TCP. A complementary overview of the applications of AI for personalized prescription is provided by the special issue edited by Rancati et al.⁶³ The studies vary from the integration of radiosensitivity biomarkers during treatment planning (similarly to Scott et al), but most of the original research studies in the collection focus on the application of AI algorithms for the modeling of toxicity outcomes in clinical cohorts. Cancer sites include brain tumors, head-and-neck, and thoracic diseases (mainly breast cancer, lung, and esophageal cancers), and prostate cancer, and substantial AI algorithm heterogeneity can be found. Some studies prefer to focus on more traditional ML algorithms such as logistic regression or random forest, up to studies that included image-derived biomarkers (radiomics/deep learning) and finally research that uses advanced AI methods such as convolutional neural networks beyond the previously cited U-NET, or even recurrent neural networks in case of time series data (*e.g.*, data acquired during each delivery fraction). It is impossible to draw a conclusion whether more traditional AI algorithms should be preferred over more complex ones, since the choice of the AI algorithms cannot be separated from the type of data that are mined. For example, in case of genomic data, more classical AI algorithms have been chosen compared to when using imaging, dosimetry or clinical data, probably showing that embedding biological knowledge of the interaction of radiation with the tumor/host helps reducing the complexity of the AI models. What emerges is the lack of studies that pool all the possible data available and

try to investigate correlations among them. Therefore, we urge the research community to collaborate to provide recommendations for the most optimal choice of the AI algorithms to be used according to the data available.

RISK-STRATIFICATION FOR OUTCOMES AND TOXICITIES

As mentioned earlier, AI algorithms can also be used to select the optimal treatment for a patient using DSSs that evaluate the likelihood of treatment response *vs* toxicities. The same DSSs can also be used to isolate groups of patients with different hazard ratios for the risk of an event such as early death, poor response, or the development of metastases. Moving from the identification of these risk groups into an actual actionable for the clinician to be discussed with the patient is not trivial. At the time of writing, many of the AI studies trying to isolate high-risk *vs* low-risk patients can be used for example to propose adjuvant treatment for the high-risk group, while preserving the low-risk group from unnecessary harmful additional radiation. Examples of these are the studies by Wu et al⁶⁴ and Mattonen et al⁶⁵ who looked at identifying early-stage lung cancer patients at higher risk of developing distant metastasis that might benefit from SBRT and/or systemic therapy instead of lobectomy, despite being in good fit for surgery. Another example of AI-based risk stratification is the development of combined imaging/clinical signatures for the identification of small cell lung cancer patients that might benefit from PCI (Prophylactic Cranial Irradiation).⁶⁶ Many studies have also associated risk stratification with treatment response prediction. These AI models can have a strong impact in supporting the selection of patients that will benefit from adjuvant treatment following RT such as immunotherapy. Examples are the studies by Duffy et al⁶⁷ and by Trebeschi et al.⁶⁸ First prospective clinical trials are currently recruiting patients to explore outcome as well as toxicity prediction by means of AI in a clinical setting, *e.g.*, for head and neck cancer (NCT05607225, NCT05081531). It is worth mentioning that almost all the studies investigating risk stratification and treatment response/ toxicity prediction made a strong use of mining imaging data. As already mentioned in the previous sections, RT planning/delivery and monitoring strongly rely on the use of medical images, which come as “cost-free” data to be analyzed. Overall, we are also seeing a trend toward using more and more complex deep learning-based algorithms to mine imaging data.

REPRODUCIBILITY, GENERALIZABILITY, AND DISTRIBUTED LEARNING

Despite the clinical endpoint investigated, we found that many AI algorithms suffer from the issues of reproducibility and generalizability. Reproducibility refers to the ability of being able to fully reproduce the results of a research study developing an AI model on another dataset, without the need of re-training the algorithm from scratch. Many studies have raised the attention over the lack of reproducibility of AI studies in RT independently from the type of algorithms chosen from the data used and for the outcome of interest.^{69,70} We can mainly distinguish two causes behind this issue: A) poor reporting and lack of standardized good practices (both for data collection and algorithms), and B) complexity of the AI models during its whole lifecycle from development to training and deployment. To solve the first issue,

consortia of researchers have been focusing on the development of guidelines that will hopefully support reproducibility of developed AI algorithms.^{71,72} Again, we have noted a lack of agreement among the different consortia, so that it is possible only to isolate a small subset of recommendations common among these publications. To solve the second issue, stating that from now on the research community should discard the use of complex AI algorithms, often borrowed from outside the medical field, is not the solution. This is because the choice of the AI algorithms and their complexity with respect to the data types that are mined is still not standardized. Nevertheless, there are methods to overcome the complexity of these algorithms for deployment and make them easily accessible to other researchers or even to clinical professionals. Examples of these are common repository such as GitHub, and modelHub.AI for release of the code, while for the straightforward use of the models, Docker containers should be the preferable solutions. Another interesting opportunity is to bypass the installation of the software (*i.e.*, an AI algorithm) within one institution and adopt a SaaS (Software as a Service) approach, which is typically achieved by Cloud infrastructures.⁷³

Next to reproducibility, the issue of generalizability of AI models, overall and for RT, remain a critical issue impeding a fast clinical translation of these powerful technologies. Generalizability refers to the capability of an AI algorithm to hold prognostic/predictive values or more in general to be robust on heterogeneous datasets, besides the data used for training. Generalizability is the opposite of overfitting. Therefore, generalizability strongly relies on the availability of a large amount of annotated data in RT. Annotating and collecting data is expensive, and annotations can vary among different RT clinics because of different clinical guidelines. If solutions like auto-contouring or auto-planning can help in reducing the burden of manually annotating data, still the issue of centrally collecting the data remains. Even in the presence of deep-learning algorithms that can generate synthetic data (*e.g.*, Generative Adversarial Models for pseudo-CT), a large dataset needs to be collected for training the algorithms. Data produced by the use of AI algorithms should be stored and re-used to re-train or fine-tune the algorithms to achieve better performances and eventually increase clinical applicability. The ecosystem produced by AI algorithms, the operators using the algorithms and the results of this interaction (*e.g.*, measured outcome *vs* predicted outcome) should move toward a self-learning healthcare system. Examples of promising applications are self-learning algorithms, or reinforcement learning.⁷⁴ Centrally transferring and collecting the data always come with concerns about privacy and the legal path to approval is much longer than the technical obstacles (almost inexistent) that need to be solved. Therefore, distributed learning as a privacy preserving solution, where the model is trained and transferred to the institutions, rather than the data, is becoming more and more popular, also supported by enterprise research on distributed learning infrastructures. Recent publications have shown not only the possibility to fully develop machine-learning algorithms in a distributed fashion for outcome prediction⁷⁵ but also for the development of distributed deep learning algorithms for auto-contouring⁷⁶ and again prognostication/predictions.⁷⁷ Performance metrics used to evaluate AI algorithms might

not be fully representative of the real clinical applicability and value of these algorithms. The literature suggests the use of DCA (Decision Curve Analysis) as an objective quantification of clinical impact. In the case of AI algorithms for segmentation, the usage of reading tests (*e.g.*, Turing test) is strongly recommended to evaluate, for example, the clinical acceptance of automated contours, compared to for example just using a DICE score. Finally, we strongly support the idea of performing clinical trials where the impact of AI algorithms in supporting clinical decisions is evaluated in a controlled setting.

ETHICAL CHALLENGES

Currently, while many AI algorithms have been proposed, few have become a clinical standard. A first requirement is that prospective evaluation is important to make sure that the results of retrospective evaluation were not biased by data selection and curation. In other words, application to real-world data is needed to determine the performance of an AI algorithm. This is important since some models rely on feature selection by the model designer (for example, radiomics models) with choices made based on the retrospective data used to train the model. This makes these models semi-automatic in nature, and potentially the selected features may not be ideal for prospective data or different populations.

In that context, ethical, medicolegal and quality control aspects are critical. Questions such as the obligation to inform patients on the use of AI tools and of responsibility when AI does harm are far from answered. The frequency of AI quality control evaluations to detect drifts in the input data is also an open question. A good summary of these points can be found in Bhowmik *et al.*⁷⁸

The interaction of humans with machines and algorithms is a field of active investigations. Introducing AI systems in the consolidated process of decision-making and RT planning/delivery requires a dedicated process of change management. Naik *et al* stated that the legal and ethical issues that confront society due to AI include privacy and surveillance, bias, or discrimination.⁷⁹ In their analysis, the authors underline how there are no well-defined regulations in place to address the legal and ethical issues that may arise due to the use of AI in a healthcare setting. Without going into detail, we believe that the key questions that we are asked to solve as the RT/oncology community is: does AI fit within the existing legal categories? Should a new category with its special features and implications be developed? Sub-questions related to the applications of AI presented in this review are: 1) "If new data are generated with AI, who is the data owner? Should informed consent be required again? 2) Since AI algorithms are trained on specific collected data, which may have human biases, is AI really supporting equality of care? 3) Is it possible to make decisions based on data that do not have real-world counterparts (*e.g.*, AI pseudo-generated/synthetic data), and 4) Are AI algorithms medical devices and do we need to re-design clinical trials to evaluate and approve AI algorithms? We believe that all these questions demand to be evaluated by not only RT and healthcare professionals but also involving policymakers and governmental agencies.

REFERENCES

- Lin T-Y, Goyal P, Girshick R, He K, Dollar P. Focal Loss for Dense Object Detection. In: Dollár P, ed. 2017 IEEE International Conference on Computer Vision (ICCV); Venice. ; 2017. <https://doi.org/10.1109/ICCV.2017.324>
- Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-Configuring method for deep learning-based BIOMEDICAL image Segmentation. *Nat Methods* 2021; **18**: 203–11. <https://doi.org/10.1038/s41592-020-01008-z>
- Marschner S, Datar M, Gaasch A, Xu Z, Grbic S, Chabin G, et al. Correction: A deep image-to-image network organ Segmentation algorithm for radiation treatment planning: principles and evaluation. *Radiat Oncol* 2022; **17**: 149. <https://doi.org/10.1186/s13014-022-02110-6>
- Harrison K, Pullen H, Welsh C, Oktay O, Alvarez-Valle J, Jena R. Machine learning for auto-Segmentation in radiotherapy planning. *Clin Oncol (R Coll Radiol)* 2022; **34**: 74–88. <https://doi.org/10.1016/j.clon.2021.12.003>
- Lin L, Dou Q, Jin Y-M, Zhou G-Q, Tang Y-Q, Chen W-L, et al. Deep learning for automated contouring of primary tumor volumes by MRI for Nasopharyngeal carcinoma. *Radiology* 2019; **291**: 677–86. <https://doi.org/10.1148/radiol.2019182012>
- Brouwer CL, Boukerroui D, Oliveira J, Looney P, Steenbakkers R, Langendijk JA, et al. Assessment of manual adjustment performed in clinical practice following deep learning contouring for head and neck organs at risk in radiotherapy. *Phys Imaging Radiat Oncol* 2020; **16**: 54–60. <https://doi.org/10.1016/j.phro.2020.10.001>
- Vaassen F, Boukerroui D, Looney P, Canters R, Verhoeven K, Peeters S, et al. Real-world analysis of manual editing of deep learning contouring in the Thorax region. *Phys Imaging Radiat Oncol* 2022; **22**: 104–10. <https://doi.org/10.1016/j.phro.2022.04.008>
- Kawula M, Hadi I, Nierer L, Vagni M, Cusumano D, Boldrini L, et al. Patient-specific transfer learning for auto-Segmentation in adaptive 0.35 T Mrgt of prostate cancer: a bi-centric evaluation. *Med Phys* 2023; **50**: 1573–85. <https://doi.org/10.1002/mp.16056>
- Wang C, Zhu X, Hong JC, Zheng D. Artificial intelligence in radiotherapy treatment planning: present and future. *Technol Cancer Res Treat* 2019; **18**: 1533033819873922. <https://doi.org/10.1177/1533033819873922>
- Hussein M, Heijmen BJM, Verellen D, Nisbet A. Automation in intensity modulated radiotherapy treatment planning-a review of recent innovations. *Br J Radiol* 2018; **91**: 20180270. <https://doi.org/10.1259/bjr.20180270>
- Moore KL. Automated radiotherapy treatment planning. *Semin Radiat Oncol* 2019; **29**: 209–18. <https://doi.org/10.1016/j.semradonc.2019.02.003>
- Nguyen D, Lin M-H, Sher D, Lu W, Jia X, Jiang S. Advances in automated treatment planning. *Semin Radiat Oncol* 2022; **32**: 343–50. <https://doi.org/10.1016/j.semradonc.2022.06.004>
- Klüter S. Technical design and concept of a 0.35 T MR-Linac. *Clinical and Translational Radiation Oncology* 2019; **18**: 98–101. <https://doi.org/10.1016/j.ctro.2019.04.007>
- Winkel D, Bol GH, Kroon PS, van Asselen B, Hackett SS, Werensteijn-Honingh AM, et al. Adaptive radiotherapy: the Elekta unity MR-Linac concept. *Clin Transl Radiat Oncol* 2019; **18**: 54–59. <https://doi.org/10.1016/j.ctro.2019.04.001>
- Corradini S, Alongi F, Andrasschke N, Belka C, Boldrini L, Cellini F, et al. MR-guidance in clinical reality: Current treatment challenges and future perspectives. *Radiat Oncol* 2019; **14**: 92. <https://doi.org/10.1186/s13014-019-1308-y>
- Kurz C, Buizza G, Landry G, Kamp F, Rabe M, Paganelli C, et al. Medical physics challenges in clinical MR-guided radiotherapy. *Radiat Oncol* 2020; **15**: 93. <https://doi.org/10.1186/s13014-020-01524-4>
- Byrne M, Archibald-Heeren B, Hu Y, Teh A, Beserminji R, Cai E, et al. Varian ethos online adaptive radiotherapy for prostate cancer: early results of contouring accuracy, treatment plan quality, and treatment time. *J Appl Clin Med Phys* 2022; **23**: e13479. <https://doi.org/10.1002/acm2.13479>
- Sahin B, Zoto Mustafayev T, Gungor G, Aydin G, Yapici B, Atalar B, et al. First 500 fractions delivered with a magnetic resonance-guided radiotherapy system: initial experience. *Cureus* 2019; **11**: e6457. <https://doi.org/10.7759/cureus.6457>
- Rogowski P, von Bestenbostel R, Walter F, Straub K, Nierer L, Kurz C, et al. Feasibility and early clinical experience of online adaptive MR-guided radiotherapy of liver tumors. *Cancers (Basel)* 2021; **13**(7): 1523. <https://doi.org/10.3390/cancers13071523>
- Hadi I, Eze C, Schönecker S, von Bestenbostel R, Rogowski P, Nierer L, et al. MR-guided SBRT boost for patients with locally advanced or recurrent gynecological cancers ineligible for Brachytherapy: feasibility and early clinical experience. *Radiat Oncol* 2022; **17**: 8. <https://doi.org/10.1186/s13014-022-01981-z>
- Lamb J, Cao M, Kishan A, Agazaryan N, Thomas DH, Shaverdian N, et al. Online adaptive radiation therapy: implementation of a new process of care. *Cureus* 2017; **9**(8): e1618. <https://doi.org/10.7759/cureus.1618>
- Guo Y, Gao Y, Shen D. Deformable MR prostate Segmentation via deep feature learning and sparse patch matching. *IEEE Trans Med Imaging* 2016; **35**: 1077–89. <https://doi.org/10.1109/TMI.2015.2508280>
- Feng Z, Nie D, Wang L, Shen D. Semi-supervised learning for pelvic Mr image Segmentation based on multi-task residual fully Convolutional networks. *Proc IEEE Int Symp Biomed Imaging* 2018; **2018**: 885–88. <https://doi.org/10.1109/ISBI.2018.8363713>
- Elgundi S, Zelefsky MJ, Jiang J, Veeraraghavan H, Deasy JO, Hunt MA, et al. Deep learning-based auto-Segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. *Phys Imaging Radiat Oncol* 2019; **12**: 80–86. <https://doi.org/10.1016/j.phro.2019.11.006>
- Kuisma A, Ranta I, Keyriläinen J, Suilamo S, Wright P, Pesola M, et al. Validation of automated magnetic resonance image Segmentation for radiation therapy planning in prostate cancer. *Phys Imaging Radiat Oncol* 2020; **13**: 14–20. <https://doi.org/10.1016/j.phro.2020.02.004>
- Gou S, Lee P, Hu P, Rwigema J-C, Sheng K. Feasibility of automated 3-dimensional magnetic resonance imaging Pancreas Segmentation. *Adv Radiat Oncol* 2016; **1**: 182–93. <https://doi.org/10.1016/j.adro.2016.05.002>
- Liang Y, Schott D, Zhang Y, Wang Z, Nasief H, Paulson E, et al. Auto-Segmentation of Pancreatic tumor in multi-parametric MRI using deep Convolutional neural networks. *Radiother Oncol* 2020; **145**: 193–200. <https://doi.org/10.1016/j.radonc.2020.01.021>
- Fu Y, Mazur TR, Wu X, Liu S, Chang X, Lu Y, et al. A novel MRI Segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. *Med Phys* 2018; **45**: 5129–37. <https://doi.org/10.1002/mp.13221>
- Chen Y, Ruan D, Xiao J, Wang L, Sun B, Saouaf R, et al. Fully automated Multiorgan Segmentation in abdominal magnetic resonance imaging with deep neural networks. *Med Phys* 2020; **47**: 4971–82. <https://doi.org/10.1002/mp.14429>
- Savenije MHF, Maspero M, Sikkes GG, van der Voort van Zyp JRN, T J Kotte AN, Bol

- GH, et al. Clinical implementation of MRI-based organs-at-risk auto-Segmentation with Convolutional networks for prostate radiotherapy. *Radiat Oncol* 2020; **15**: 104. <https://doi.org/10.1186/s13014-020-01528-0>
31. Gurney-Champion OJ, Kieselmann JP, Wong KH, Ng-Cheng-Hin B, Harrington K, Oelfke U. A Convolutional neural network for contouring metastatic lymph nodes on diffusion-weighted magnetic resonance images for assessment of radiotherapy response. *Phys Imaging Radiat Oncol* 2020; **15**: 1–7. <https://doi.org/10.1016/j.phro.2020.06.002>
 32. Cusumano D, Boldrini L, Dhont J, Fiorino C, Green O, Güngör G, et al. Artificial intelligence in magnetic resonance guided radiotherapy: medical and physical considerations on state of art and future perspectives. *Phys Med* 2021; **85**: 175–91. <https://doi.org/10.1016/j.ejmp.2021.05.010>
 33. Chun J, Park JC, Olberg S, Zhang Y, Nguyen D, Wang J, et al. Intentional deep Overfit learning (IDOL): A novel deep learning strategy for adaptive radiation therapy. *Med Phys* 2022; **49**: 488–96. <https://doi.org/10.1002/mp.15352>
 34. Chen X, Ma X, Yan X, Luo F, Yang S, Wang Z, et al. Personalized auto-Segmentation for magnetic resonance imaging-guided adaptive radiotherapy of prostate cancer. *Med Phys* 2022; **49**: 4971–79. <https://doi.org/10.1002/mp.15793>
 35. Li Z, Zhang W, Li B, Zhu J, Peng Y, Li C, et al. Patient-specific daily updated deep learning auto-Segmentation for MRI-guided adaptive radiotherapy. *Radiother Oncol* 2022; **177**: 222–30. <https://doi.org/10.1016/j.radonc.2022.11.004>
 36. Fransson S, Tilly D, Strand R. Patient specific deep learning based Segmentation for magnetic resonance guided prostate radiotherapy. *Phys Imaging Radiat Oncol* 2022; **23**: 38–42. <https://doi.org/10.1016/j.phro.2022.06.001>
 37. Eppenhof KAJ, Maspero M, Savenije MHF, de Boer JCJ, van der Voort van Zyp JRN, Raaymakers BW, et al. Fast contour propagation for MR-guided prostate radiotherapy using Convolutional neural networks. *Med Phys* 2020; **47**: 1238–48. <https://doi.org/10.1002/mp.13994>
 38. Brion E, Léger J, Barragán-Montero AM, Meert N, Lee JA, Macq B. Domain adversarial networks and intensity-based data augmentation for male pelvic organ Segmentation in cone beam CT. *Comput Biol Med* 2021; **131**: 104269. <https://doi.org/10.1016/j.compbiomed.2021.104269>
 39. Spadea MF, Maspero M, Zaffino P, Seco J. Deep learning based synthetic-CT generation in radiotherapy and PET: A review. *Med Phys* 2021; **48**: 6537–66. <https://doi.org/10.1002/mp.15150>
 40. Gurney-Champion OJ, Landry G, Redalen KR, Thorwarth D. Potential of deep learning in quantitative magnetic resonance imaging for personalized radiotherapy. *Semin Radiat Oncol* 2022; **32**: 377–88. <https://doi.org/10.1016/j.semradonc.2022.06.007>
 41. Han X. MR-Based synthetic CT generation using a deep Convolutional neural network method. *Med Phys* 2017; **44**: 1408–19. <https://doi.org/10.1002/mp.12155>
 42. Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku J, et al. Cone beam computed tomography image quality improvement using a deep Convolutional neural network. *Cureus* 2018; **10**: e2548. <https://doi.org/10.7759/cureus.2548>
 43. Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q. Medical image synthesis with context-aware generative adversarial networks. International conference on medical image computing and computer-assisted intervention. Springer; 2017. <https://doi.org/10.1007/978-3-319-66179-7>
 44. Kurz C, Maspero M, Savenije MHF, Landry G, Kamp F, Pinto M, et al. CBCT correction using a cycle-consistent Generative adversarial network and unpaired training to enable photon and proton dose calculation. *Phys Med Biol* 2019; **64**(22): ab4d8c. <https://doi.org/10.1088/1361-6560/ab4d8c>
 45. Cai J, Zhang Z, Cui L, Zheng Y, Yang L. Towards cross-modal organ translation and Segmentation: A Cycle- and shape-consistent Generative adversarial network. *Med Image Anal* 2019; **52**: 174–84. <https://doi.org/10.1016/j.media.2018.12.002>
 46. Yang H, Sun J, Carass A, Zhao C, Lee J, Xu Z. Unpaired Brain MR-to-CT Synthesis Using a Structure-Constrained CycleGAN. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer International Publishing; 2018. <https://doi.org/10.1007/978-3-030-00889-5>
 47. Eze C, Lombardo E, Nierer L, Xiong Y, Niyazi M, Belca C, et al. MR-guided radiotherapy in node-positive non-small cell lung cancer and severely limited pulmonary Reserve: a report proposing a new clinical pathway for the management of high-risk patients. *Radiat Oncol* 2022; **17**: 43. <https://doi.org/10.1186/s13014-022-02011-8>
 48. Keall PJ, Sawant A, Berbeco RI, Booth JT, Cho B, Cervino LI, et al. AAPM task group 264: the safe clinical implementation of MLC tracking in radiotherapy. *Med Phys* 2021; **48**: e44–64. <https://doi.org/10.1002/mp.14625>
 49. Uijtewaal P, Borman PTS, Woodhead PL, Hackett SL, Raaymakers BW, Fast MF. Dosimetric evaluation of MRI-guided multi-leaf Collimator tracking and trailing for lung stereotactic body radiation therapy. *Med Phys* 2021; **48**: 1520–32. <https://doi.org/10.1002/mp.14772>
 50. Uijtewaal P, Borman PTS, Woodhead PL, Kontaxis C, Hackett SL, Verhoeff J, et al. First experimental demonstration of VMAT combined with MLC tracking for single and multi fraction lung SBRT on an MR-Linac. *Radiother Oncol* 2022; **174**: S0167-8140(22)04188-3: 149–57. <https://doi.org/10.1016/j.radonc.2022.07.004>
 51. Liu PZY, Dong B, Nguyen DT, Ge Y, Hewson EA, Waddington DEJ, et al. First experimental investigation of simultaneously tracking two independently moving targets on an MRI-Linac using real-time MRI and MLC tracking. *Med Phys* 2020; **47**: 6440–49. <https://doi.org/10.1002/mp.14536>
 52. Friedrich F, Hörner-Rieber J, Renkamp CK, Klüter S, Bachert P, Ladd ME, et al. Stability of conventional and machine learning-based tumor auto-Segmentation techniques using Undersampled dynamic radial bSSFP acquisitions on a 0.35 T hybrid MR-Linac system. *Med Phys* 2021; **48**: 587–96. <https://doi.org/10.1002/mp.14659>
 53. Glitzner M, Woodhead PL, Borman PTS, Lagendijk JJW, Raaymakers BW. Technical NOTE: MLC-tracking performance on the Elekta unity MRI-Linac. *Phys Med Biol* 2019; **64**: 15NT02. <https://doi.org/10.1088/1361-6560/ab2667>
 54. Yun J, Wachowicz K, Mackenzie M, Rathee S, Robinson D, Fallone BG. First demonstration of Intrafractional tumor-tracked irradiation using 2d phantom MR images on a prototype Linac-MR. *Med Phys* 2013; **40**(5): 051718. <https://doi.org/10.1118/1.4802735>
 55. Lombardo E, Rabe M, Xiong Y, Nierer L, Cusumano D, Placidi L, et al. Offline and online LSTM networks for respiratory motion prediction in MR-guided radiotherapy. *Phys Med Biol* 2022; **67**(9). <https://doi.org/10.1088/1361-6560/ac60b7>
 56. Lombardo E, Rabe M, Xiong Y, Nierer L, Cusumano D, Placidi L, et al. Evaluation of real-time tumor contour prediction using LSTM networks for MR-guided radiotherapy. *Radiother Oncol* 2023; **182**: S0167-8140(23)00093-2: 109555. <https://doi.org/10.1016/j.radonc.2023.109555>
 57. Hunt B, Gill GS, Alexander DA, Streeter SS, Gladstone DJ, Russo GA, et al. Fast Deformable image registration for real-time target tracking during radiation therapy using cine MRI and deep learning. *Int J Radiat Oncol Biol Phys* 2023; **115**: S0360-

- 3016(22)03444-7: 983–93: . <https://doi.org/10.1016/j.ijrobp.2022.09.086>
58. Galmarini CM. Why we do what we do. A brief analysis of cancer therapies. *EXCLI J* 2020; **19**: 1401–13. <https://doi.org/10.17179/excli2020-2972>
 59. Ho D, Quake SR, McCabe ERB, Chng WJ, Chow EK, Ding X, et al. Enabling Technologies for personalized and precision medicine. *Trends in Biotechnology* 2020; **38**: 497–518. <https://doi.org/10.1016/j.tibtech.2019.12.021>
 60. Tang M, Pearson SA, Simes RJ, Chua BH. Harnessing real-world evidence to advance cancer research. *Curr Oncol* 2023; **30**: 1844–59. <https://doi.org/10.3390/curroncol30020143>
 61. Hak F, Guimarães T, Santos M. Towards effective clinical decision support systems: A systematic review. *PLoS One* 2022; **17**(8): e0272846. <https://doi.org/10.1371/journal.pone.0272846>
 62. Scott JG, Sedor G, Scarborough JA, Kattan MW, Peacock J, Grass GD, et al. Personalizing radiotherapy prescription dose using Genomic markers of Radiosensitivity and normal tissue toxicity in NSCLC. *J Thorac Oncol* 2021; **16**: S1556-0864(20)31046-7: 428–38: . <https://doi.org/10.1016/j.jtho.2020.11.008>
 63. Rancati T, Fiorino C, Sanguineti G, Valdagni R, Orlandi E. Editorial: modeling for prediction of radiation-induced toxicity to improve therapeutic ratio in the modern radiation therapy era. *Front Oncol* 2021; **11**: 690649. <https://doi.org/10.3389/fonc.2021.690649>
 64. Wu J, Aguilera T, Shultz D, Gudur M, Rubin DL, Loo BW Jr, et al. Early-stage non-small cell lung cancer: quantitative imaging characteristics of (18)F Fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology* 2016; **281**: 270–78. <https://doi.org/10.1148/radiol.2016151829>
 65. Mattonen SA, Davidzon GA, Benson J, Leung ANC, Vasanawala M, Horng G, et al. Bone Marrow and Tumor Radiomics at (18)F-FDG PET/CT: Impact on Outcome Prediction in Non-Small Cell Lung Cancer. *Radiology*. 2019;293(2):451–59.
 66. Zeng H, Zheng D, Witlox W, Levy A, Traverso A, Kong F-MS, et al. Ep14.01-014 risk factors for brain metastasis in patients with small cell lung cancer: A systematic review and meta-analysis. *Journal of Thoracic Oncology* 2022; **17**: S532. <https://doi.org/10.1016/j.jtho.2022.07.950>
 67. Duffy MJ, Crown J. Biomarkers for predicting response to Immunotherapy with immune Checkpoint inhibitors in cancer patients. *Clin Chem* 2019; **65**: 1228–38. <https://doi.org/10.1373/clinchem.2019.303644>
 68. Trebeschi S, Drago SG, Birkbak NJ, Kurilova I, Călin AM, Delli Pizzi A, et al. Predicting response to cancer Immunotherapy using noninvasive Radiomic biomarkers. *Ann Oncol* 2019; **30**: S0923-7534(19)31202-5: 998–1004: . <https://doi.org/10.1093/annonc/mdz108>
 69. Pfähler E, Zhovannik I, Wei L, Boellaard R, Dekker A, Monshouwer R, et al. A systematic review and quality of reporting checklist for Repeatability and reproducibility of Radiomic features. *Phys Imaging Radiat Oncol* 2021; **20**: 69–75. <https://doi.org/10.1016/j.phro.2021.10.007>
 70. Gundersen OE, Kjensmo S. State of the art: reproducibility in artificial intelligence. *AAAI* 2018; **32**(1). <https://doi.org/10.1609/aaai.v32i1.11503>
 71. Matschinske J, Alcaraz N, Benis A, Golebiewski M, Grimm DG, Heumos L, et al. The Aime Registry for artificial intelligence in BIOMEDICAL research. *Nat Methods* 2021; **18**: 1128–31. <https://doi.org/10.1038/s41592-021-01241-0>
 72. Sanduleanu S, Woodruff HC, de Jong EEC, van Timmeren JE, Jochems A, Dubois L, et al. Tracking tumor biology with Radiomics: A systematic review utilizing a Radiomics quality score. *Radiother Oncol* 2018; **127**: S0167-8140(18)30179-8: 349–60: . <https://doi.org/10.1016/j.radonc.2018.03.033>
 73. Turner M, Budgen D, Brereton P. Turning software into a service. *Computer* 2003; **36**: 38–44. <https://doi.org/10.1109/MC.2003.1236470>
 74. Krishnamurthy R, Mummudi N, Goda JS, Chopra S, Heijmen B, Swamidas J. Using artificial intelligence for optimization of the processes and resource utilization in radiotherapy. *JCO Glob Oncol* 2022; **8**: e2100393. <https://doi.org/10.1200/GO.21.00393>
 75. Shi Z, Zhovannik I, Traverso A, Dankers FJWM, Deist TM, Kalendralis P, et al. Distributed Radiomics as a signature validation study using the personal health train infrastructure. *Sci Data* 2019; **6**(1): 218. <https://doi.org/10.1038/s41597-019-0241-0>
 76. Remedios SW, Roy S, Bermudez C, Patel MB, Butman JA, Landman BA, et al. Distributed deep learning across Multisite Datasets for generalized CT hemorrhage Segmentation. *Med Phys* 2020; **47**: 89–98. <https://doi.org/10.1002/mp.13880>
 77. Zerka F, Barakat S, Walsh S, Bogowicz M, Leijenaar RTH, Jochems A, et al. Systematic review of privacy-preserving distributed machine learning from Federated databases in health care. *JCO Clin Cancer Inform* 2020; **4**: CCI.19.00047: 184–200: . <https://doi.org/10.1200/CCI.19.00047>
 78. Bhowmik A, Eskreis-Winkler S. Deep learning in breast imaging. *BJR Open* 2022; **4**(1): 20210060. <https://doi.org/10.1259/bjro.20210060>
 79. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, et al. Legal and ethical consideration in artificial intelligence in Healthcare: who takes responsibility *Front Surg* 2022; **9**: 862322. <https://doi.org/10.3389/fsurg.2022.862322>