The Breast 83 (2025) 104537

Contents lists available at ScienceDirect

The Breast

journal homepage: www.journals.elsevier.com/the-breast

Artificial intelligence in breast cancer radiotherapy: Insights from the Toolbox Consortium Delphi study

Orit Kaidar-Person [a,b,*], André Pfob [c,d], Vincenzo Valentini [e], Marianne Aznar [f], Andre Dekker [g], Icro Meattini [h], Jana de Boniface [i,j], David Krug [k], Maria Joao Cardoso [l], Giuseppe Curigliano [m,n], Peter Dubsky [o,p], Philip Poortmans [q,r]

[a] Breast Radiation Unit, Sheba Tel Hashomer, Ramat Gan, Israel
[b] Gray School of Medical Sciences, Tel-Aviv University, Tel-Aviv, Israel
[c] Breast Center Heidelberg, Hospital St. Elisabeth, Heidelberg, Germany
[d] Department of Obstetrics & Gynecology, Heidelberg University Hospital, Heidelberg, Germany
[e] Centro Eccellenza Oncologia e Diagnostica per Immagini, Ospedale Isola Tiberina – Gemelli Isola, Roma, Italy
[f] Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[g] Department of Radiation Oncology (Maastro), GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, the Netherlands
[h] Radiation Oncology Unit, Azienda Ospedaliero Universitaria Careggi-University of Florence, Florence, Italy
[i] Department of Surgery, Capio St. Göran's Hospital, Stockholm, Sweden
[j] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[k] Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg-Eppendorf AND Department of Radiation Oncology, University Medical Center Schleswig-Holstein, Kiel, Germany
[l] Breast Cancer Research Program, Champalimaud Foundation and Lisbon University Faculty of Medicine, Portugal
[m] European Institute of Oncology, IRCCS, Milano, Italy
[n] University of Milano, Milano, Italy
[o] Breast Centre, Hirslanden Klinik St. Anna, Luzern, Switzerland
[p] University of Lucerne, Department of Health Sciences and Medicine, Luzern, Switzerland
[q] Department of Radiation Oncology, Iridium Netwerk, Wilrijk-Antwerp, Belgium
[r] University of Antwerp, Faculty of Medicine and Health Sciences, Wilrijk-Antwerp, Belgium

A B S T R A C T

Artificial intelligence (AI) is being incorporated in several breast cancer care domains, including for radiation therapy (RT). Herein we provide a review about AI for the management and planning of RT for breast cancer, which is part of the Toolbox-3 project's multidisciplinary Delphi study, including a literature review of studies related to the topic raised by the Delphi questionnaire. Our review shows that available evidence mainly consists of small single institutional studies, often at least partly supported by commercial companies. Current studies suffer from a lack of transparency regarding how these systems were developed, the information they are based on, the algorithms used, and potential proprietary issues. This review provides a critical inter- and multidisciplinary assessment of existing systems to help us in guiding development and utilisation of AI-based tools in the field of radiation oncology. As medical professional users, we must remain vigilant and continue to improve our personal experience and knowledge that serves as the "ground truth". Employing AI required a critical mindset, particularly in medical applications which may influence the lives of our patients.

1. Introduction

The Toolbox Consortium organises a bi-annual meeting held in Lucerne, Switzerland, focusing on providing practical clinical recommendations for challenging or evidence-lacking dilemmas in daily breast cancer clinical practice [1,2]. An international, multidisciplinary/interdisciplinary team identifies the topic of the meeting and forms working groups (WGs) according to expertise, to identify key issues for discussion that are later consolidated into working packages (WPs). Preparatory work includes reviewing the literature and designing a questionnaire to facilitate voting as part of a modified Delphi process [1–3]. A pre-meeting vote is conducted to determine the final

* Corresponding author. Breast Radiation Unit, Sheba Tel Hashomer, Ramat Gan, Israel. Gray School of Medical Sciences, Tel Aviv University, Tel Aviv, Israel
E-mail address: orit.kaidarperson@sheba.health.gov.il (O. Kaidar-Person).

[1] The list of the toolbox consortium members/coauthors are listed in appendix

https://doi.org/10.1016/j.breast.2025.104537
Received 9 May 2025; Received in revised form 8 June 2025; Accepted 13 July 2025
Available online 15 July 2025
0960-9776/

key issues to be discussed at the meeting, culminating in the Toolbox recommendations. This process is designed in close collaboration with patient advocates, ensuring the relevance of the recommendations for the clinical practice of breast cancer care from the patients' perspective.

With the increasing prominence of Artificial intelligence (AI) in healthcare and research, the Toolbox-3 aimed to provide guidance on priority gaps for the implementation of AI tools in the treatment of breast cancer. Although discussed and voted the questions on radiation oncology were not discussed at the face-to-face meeting according to the priorities established for discussion.

AI is increasingly being incorporated into various aspects of radiation oncology technology and clinical practice. Recognising the urgent need for guidance, radiation therapy (RT) was the first medical technology to have a National Institute for Health and Care Excellence (NICE)-approved AI-based recommendation for cancer care [4]. Major issues that were identified by the radiation interdisciplinary team as important for discussion in the Toolbox framework were: AI-based decision support for radiation indication, dose and volume, prediction of radiation-related toxicity, mainly cardiac toxicity, lymphedema, and aesthetic outcomes, and AI for real-time treatment adaptions including matching, breast shape/tissue changes, early signs of RT-related side effects, and improving quality assurance measures. However, based on the pre-meeting votes, these topics were not selected for discussion at the Toolbox-3 face to face meeting as they did not receive sufficient priority to warrant further discussion (the score is presented in Table 1, proportion of recommendations for research agenda). This narrative review aims to evaluate whether their low-priority grading reflects a "vote of confidence" in the established AI applications in radiation oncology, or perhaps "overconfidence" [5]. Toolbox recommendations presented in this work is based on discussions among radiation oncology team and critical review of the summary by the toolbox members.

2. Methods

2.1. Lucerne toolbox procedures

In brief, the steering committee and expert panel selection were chosen to assure an international representation, linked to international/national societies of multidisciplinary expertise, including patient advocates [1–3]. The steering committee designed six WGs to develop WPs according to each speciality. The full description of the Toolbox procedure is available in the previous publications [1,3], and is planned to be published with the main Toolbox-3 results. The pre-meeting procedures aimed to identify the 15 most important knowledge gaps to be discussed at the consensus conference.

The Delphi survey was sent to 154 participants on July 25th, 2024, and contained a list of potential knowledge gaps related to digital health tools applicable throughout the patient journey for women with early breast cancer. Participants were asked to rank the importance of every knowledge gap on a 9-point Likert scale from 1 (not important) to 9 (extremely important). Additionally, they were asked to select a maximum of 10 topics to be discussed at the consensus conference. To determine the 15 most important knowledge gaps to be discussed, the top 10 knowledge gaps from the Likert ranking as well as the top 10 knowledge gaps recommended by the experts were selected. To consider potential imbalances between disciplines, the Likert ratings and proportion of recommended knowledge gaps were adjusted to the number of participants per discipline (i.e. disciplines with fewer participants were given more weight).

The radiation oncology WG was composed of six international experts, whose work included email discussions and a teleconference for selecting radiation-related questions to be voted upon in the first round.

Nine out of 54 proposed questions were related to radiation oncology. These questions were not voted as priority questions but the Delphi results were recorded and discussed among the steering committee and radiotherapy experts.

Table 1
Radiation oncology questions sent to the participants and their mean Likert rating.

Question # in the survey, the question received by participants	Mean Likert rating	Proportion of recommendations for research agenda
Target volume delineation		
33. AI-based automatic delineation is already routinely used in some center. Does AI-based automatic delineation have a sufficient health technology assessment in radiation oncology? (If you think the technology assessment is not sufficient, you should rate this question as high/important)	6.46	9.09 %
34. The identification of the target volume can differ substantially after different types of surgery. What is the role of AI-based tools for the detection of the target volume in breast cancer radiation oncology after breast-conserving surgery, mastectomy, oncoplastic surgery, and for nodal irradiation?	**6.96**	11.22 %
35. Early research indicates that AI may help in the identification of residual breast or lymph node tissue after surgery. What is the development priority for AI-based tools to evaluate residual breast/lymph node tissue after mastectomy/ALND?	**6.95**	7.05 %
Treatment planning		
36. Early research indicates that AI-based integromic approaches may help individualizing radiation therapy. How important is the investigation of AI-based, individualized integromic approaches (i.e. adaption of radiation oncology according to germline mutations, circulating tumor cells/DNA)	6.87	4.75 %
37. Not every patient may need the same dosage or fractionation depending on tumor size, tumor biology etc. What is the development priority for digital decision support tools for individualized treatment planning with respect to dosage and fractionation?	**7.13**	19.76 %
38. What is the role of digital decision support tools for individualized treatment selection with respect to indication for EBRT, adaptive RT, dose-guided RT, proton therapy, brachy, MR-linac?	6.48	4.26 %
Evaluation of treatment plans		
39. Prediction models exist to demonstrate the risk radiation therapy induced normal tissue complications as well as for the tumor control probability. Have AI-based models a preferred clinical utility, validity, and efficacy over NTCP and TCP models in breast cancer?	6.57	8.10 %
40. Early research indicates that AI may help in prediction potential side effects. What is the development priority for AI-based risk assessment for potential side	6.75	2.62 %

(continued on next page)

Table 1 (*continued*)

Question # in the survey, the question received by participants		Mean Likert rating	Proportion of recommendations for research agenda
	effects? E.g., inhomogeneity (inframammary fold, nipple)		
41.	Early research indicates that AI may help in identifying areas that have not received a sufficient radiation dose. How important is the investigation of AI-based models to identify areas that are underdosed?	6.85	9.62 %

AI-artificial intelligence; RT-radiation therapy; ALND-axillary lymph node dissection; NTCP- conventional normal tissue complication probability; TCP-tumour control probability; EBRT- External Beam Radiation Therapy. In bold font: 3 highest scores.

3. Results

A total of 109 panel members replied to the survey. Table 1 presents the radiation oncology-related delphi questions that were sent to the participants and their mean Likert rating, in bold font are the 3 highest scores. The Toolbox-3 Delphi (not including RT discussion) will be available in a separate publication.

Among radiation oncology-related questions, the highest scores were granted to individualized treatment planning with respect to dosage and fractionation for each case and to the individualization of target volumes (Table 1).

4. Discussion

The Toolbox-3 aimed to provide guidance on the implementation of digital/AI tools in the treatment of breast cancer, including RT. The three topics that were selected by the radiation WG included target volume delineation, radiation planning and evaluation of planning (quality assurance in RT) which are pillars of daily clinical practice in radiation oncology.

RT as a treatment modality relies on novel technology and requires high accuracy in planning and delivery of therapy and is a medical field that can significantly benefit from AI. Breast cancer, especially in the curative setting (postoperative RT) is estimated to be 30–40 % of the workload in radiation oncology units [6]. Therefore, systems to assist in delineation, planning, and quality assurance might significantly reduce treatment load and improve patient care. The categories that were selected by the RT WG include all aspects of RT. These are listed below with. a short discussion by the WG and Toolbox recommendations.

4.1. Delineation of target volumes [Questions 33–35]

The mean Likert score for Question 33 was the second lowest among all radiation-related questions, which may indicate a "vote of confidence" in the current AI-based automatic delineation systems used for target volume delineation in clinical practice. In contrast, Questions 34 and 35 ranked among the top three, both highlighting the interest in tools which assist with identifying specific high-risk regions for target volume delineation, such as residual tissue that might harbour cancer cells.

The WG acknowledges that automatic delineation is widely available and routinely utilised to assist radiation oncologists in defining the clinical target volumes of the breast, chest wall, and regional lymphatics. AI models for automated delineation of target volumes are developed using varying methodologies. Some systems are built on large datasets of delineations ("prior knowledge"), while others employ statistical models of shape and appearance to different extents, integrating automated segmentation based on machine learning. Atlas-based segmentation solutions, however, often require adjustments in small volume delineations [7].

Several AI-based systems for automated delineation are commercially available, including models based on deep learning. Additionally, some institutions utilise in-house developed systems to assist in the delineation of organs at risk (OARs) and target volumes. These systems exhibit varying degrees of performance, as evaluated by the "correctness" of delineated organs or target volumes, the extent of edits required, and the time consumed [8–13]. In addition, an AI-based segmentation system for intraoperative radiation can create synthetic CT images from segmented cone beam CT data for delineation of the target, providing a more accurate dose calculation and optimising the adaptive 3D planning. This was suggested to enable a more precise intraoperative treatment within the limited time frame of breast surgery [14].

Table 2 summarizes examples of AI-based systems available for auto-delineation and their reported performance.

Choi et al., [15] compared the feasibility of AI-based auto-segmentation (DLBAS) (deep learning based tool for segmentation) with two commercially available atlas-based segmentation solutions: Mirada's Workflow Box (WFB, Mirada Medical, Ltd., Oxford, UK) and MIM Maestro (MIM Software Inc., Cleveland, OH) for breast cancer RT, both are non-AI digital tools. Comparisons were based on planning CT scans with intravenous contrast, which is not commonly indicated or used for breast RT, even when regional nodes are irradiated [16]. The DLBAS produced consistent results, achieving the highest average Dice similarity coefficient values and the lowest Hausdorff metric particularly for target volumes and cardiac substructures. In contrast, the ABAS performed poorly in soft tissue-based regions, such as the oesophagus, cardiac arteries, and smaller target volumes [15]. It should be noted that these systems are subject to ongoing upgrades, and it is possible that their performance outcomes have improved significantly since this publication [15]. RaySearch (RaySearch laboratories, AB, Sweden) developed a deep learning model framework for delineating breast and regional node target volumes (levels 1–4), which included OARs that were clinically evaluated [12,17]. Bakx and colleagues [17,18] evaluated the RaySearch autodelineation in comparison to manual delineation. Their analysis included the need for edits, overall delineation time, and a quantitative scoring using the Dice similarity coefficient, comparing both the automated (non-corrected) and the automated (corrected) delineations with manual delineations. The target volumes did not differ significantly between the system's outputs and the experts' delineations, with 92 % of the auto-delineations being scored as clinically acceptable, with or without corrections. The total time required for delineation was significantly reduced when using the automated system for both OARs and target volumes [17,18]. The authors concluded that this deep learning auto-delineation system can produce acceptable contours for breast volumes and regional nodes [17,18]. These findings were consistent with Almberg's evaluation of the RaySearch auto-delineation system [12]. Meixner et al., [19] evaluated three commercially available auto segmentation models for the delineation of breast cancer target volumes including [1]: breast/chest wall, different axillary levels and internal mammary chain. The models evaluated were the deep learning algorithm software of the RayStation treatment planning system (TPS, version 11B, RSL Breast CT 11B (v1.0.0.1), RaySearch Laboratories), local server and deep learning-based auto-segmentation software Limbus (Limbus Contour, Limbus Contour 1.8.0-B2, Limbus AI Inc., Regina, SK, Canada), and guideline based deep learning auto-segmentation software MVision (MVision AI, Version 1.2.2, Helsinki, Finland). The delineations were compared to reference volumes and the authors concluded that all three AI-models for auto-segmentation programs showed high-quality accuracy and provided standardization for guideline-based target volume contouring [19].

The Radiation Planning Assistant (RPA) is a web-based system capable of performing automated delineation of target volumes and OARs on planning CT scans, using "prior knowledge" via a delineation

Table 2
AI-based systems available for breast cancer auto-delineation of target volumes.

First author [Publication year]	System	Indication	Dataset cases	Volume of interest delineated	Subjective clinical evaluation (Need for corrections)	Contouring time saved with AI assistance	Quantitative evaluation	
Liu [2021]	RTD-Net	Post mastectomy	99	1 CTV (Breast)	99.3 %/98.9 % No or minor 0.7 %/1.1 % major[a]	50 %	Dice similarity coefficient and Hausdorff distance	0.9 5.65 mm
Almberg [2022]	RaySearch	Left breast whole breast Regional lymphatics	200	7 CTVs (Breast, L1-4, IMN, interpec) 11 OARs	14 % no 71 % minor 15 % major[b] 72 % no 26 % minor 2 % major	75 %		Only specific score per each volume reported
Liu [2021]	U-ResNet	Whole breast	160	1 CTV (Breast) 4 OARs	99.4 % no or minor 0.6 % major 95 % no 5 % major[a]	60 %		0.94 4.31 mm

AI- Artificial intelligence; CTV- clinical target volume; L1-4 – regional lymph node level 1–4, IMN- internal mammary node; OARs- organs at risk.
[a] According to 2 Radiation Oncologists.
[b] According to 5 Radiation Oncologists.

atlas. Additionally, the RPA can generate RT plans using an AI-based system [20,21]. Target volumes encompass different cancer types and various indications for RT. In the context of breast cancer RT, planning for postmastectomy RT is described; however, it remains unclear from the publications whether the automatic delineation incorporates AI components or relies solely on a delineation atlas. Furthermore, it is unspecified which breast cancer-related target volumes are available (e.g., nodal volumes, reconstructed breast, or chest wall). Additionally, there is a lack of clarity regarding the use of reference guidelines for target volumes, such as those provided by the Radiation Therapy Oncology Group (RTOG) or the European Society for Radiotherapy and Oncology (ESTRO), which differ significantly [16,20]. Court et al. [20] stated that delineation is of the chest wall, “target”, and OARs, but not whether it is feasible to plan for postmastectomy RT in case of reconstruction and which further target volumes are available. The authors indicate that for breast RT, the plan is based on tangential fields with a “matching supraclavicular field”; thus, it appears from the publication that the nodal target volumes are rather field-based and the system delineates levels 3–4 only [20], which can lead to underdosing or not covering high-risk nodal volumes that were not dissected [16].

The Danish Breast Cancer Group (DBCG) developed a non-commercial deep learning auto-segmentation model to delineate the lymph nodes target volumes in high-risk breast cancer patients for national clinical use [22]. The “ground truth” delineations were based on a consensus dataset. The models were trained on multicentre representative consensus data to reduce bias and reflect variability of data. Later, the models were evaluated against the interobserver variation (IOV) between breast cancer experts in delineation in a separate gold standard dataset and evaluated qualitatively by a national board. This deep learning auto-segmentation was found to performed within the IOV of an expert group and achieved comparable clinical acceptance scores to the manual delineations. This process reflects the meticulous process needed to achieve a high-performance model for auto-segmentation involving high quality data, peer review and accounting for IOV [22].

Toolbox recommendations: All stakeholders must be aware of the potential harm in blindly adopting these systems (e.g., error in auto-delineation and overtreatment and exposure of OAR or underdosing of the target). All automatic delineation systems require review and revision by an expert before being used for therapy planning. Consequently, it is essential that radiation oncologists in training are taught to manually delineate volumes based on current anatomical atlases and clinical considerations, maintaining their position as the "ground truth" and ensuring a critical evaluation of the segmentation provided by the AI system. The use of Digital/AI help in delineation will be very useful as a co-pilot however human supervision needs to be present avoiding poorer outcomes, rather than improving them [4].

4.2. Treatment planning [Questions 36–38]

The mean Likert score for Question 38 was scored low priority, which may indicate a "vote of confidence" in the ability of the radiation team or the AI to select the appropriate technique for irradiation. In contrast, Questions 36 and 37 received high scores, with Question 37 achieving the highest score among all radiation-related questions. These findings highlight the need for assistance in identifying and ensuring adequate dose coverage and dose adaptation based on the risk of residual tumour cells and the intrinsic properties of the malignancy.

AI is integrated into RT planning including advanced adaptive systems at various levels, to reduce the significant workload associated with re-planning in the event of major anatomical changes at the time of treatment [5]. Table 3 summarizes examples of AI-based systems for RT planning that are available. Since it is impossible to keep track of all the updates implemented in different systems used for AI, the discussion below includes key publications that are related to breast cancer RT planning.

The RPA system optimizes dosage distributions optimisation by adding field-in-field segments (the AI elements are not clear based on the literature). Initial testing has shown that physicians accept around 50 % of the plans as suggested and require minor changes for the remaining plans [20]. Plan optimisation and dose calculation is automatically performed in the Eclipse Treatment Planning System (Varian Medical Systems) with dose verification performed using Mobius (Varian Medical Systems). Additional AI-based systems for RT plan optimisation are now commercially available. The EZFluence (Radformation, New York, NY) in Eclipse (Varian Medical Systems VMS, Palo Alto, CA) is one of several major systems commercially available for optimising breast RT plans. The EZFluence is an automated 3D planning software that automatically generates an optimal fluence file that is used to construct the tangential based field-in-field plan. The plan may be directly imported into Varian’s Eclipse treatment planning software [23]. Compared to manual planning, the EZFluence-based planning showed better homogeneity in breast RT without significant improvement in OARs exposure but with a significant reduction of treatment planning time [23].

The votes signify the need for incorporating AI into decisions of dose, fractionation, and planning according to the different target volume. For breast cancer planning, often the clinical target volume is the breast or chest wall, with/without regional lymphatics [16], while the tumour bed or potentially gross residual disease within the nodes (non-operated axilla or internal mammary nodes) are considered as high-risk regions that should be considered for a higher radiation dose [24,25]. Breast

Table 3
Current AI-based systems available for breast cancer RT planning.

First author [Publication year]	System	Ability	Plan optimisation	Time	Comments
Court [2023]	RPA Compared to manual	Auto-delineation Planning	Tangential Field-in-field segments Supra field[a]	12 min time to edit autogenerated plan	accept around 50 %
Bakx [2023] van de Sande [2021]	RS Compared to 2 non-AI autoseg	Auto-delineation Planning	Tangential Field-in-field segments	Reduced	
Yoder [2019]	EZFluence[b]	Planning	Tangential Field-in-field segments	Reduced	Better dose optimisation Similar OAR dose

RPA -Radiation Planning Assistant; RS – RaySearch; OAR- Organs at risk.
[a] As indicated in the publication.
[b] Varian Medical Systems.

cancer RT doses are generic and evolved over decades mainly due to UK trials based also on radiobiology calculations, tumour control probability and normal tissue toxicity [25–27].

A genomic-adjusted radiation dose (GARD) is currently under development, aiming to adapt radiation dosing based on biological effect rather than relying solely on physical dose. This approach builds on preclinical work identifying tumour molecular signatures that correlate with radiosensitivity. However, the clinical utility of GARD remains to be demonstrated. Its integration as an assay in validating clinical trials will be essential to drive practice change. International collaboration and integrating such tools in future clinical trials will allow to incorporate these tools in daily practice to guide RT planning beyond current applications.

Toolbox recommendations: Current treatment planning systems can assist in RT planning and enhance both the workflow and workload management within the radiation department. It is essential for the interdisciplinary team to be adequately trained to use these applications efficiently while possessing sufficient knowledge to critically evaluate the system's outputs. Initiatives such as the application of the GARD to predict tumour control probability (TCP) and normal tissue complication probability (NTCP) on an individual-patient basis should be actively supported by grant agencies and the research community. Such projects are crucial for enabling the future personalisation of RT based on biological factors [28].

4.3. Evaluation of treatment plans

The mean Likert scores for Questions 39 and 40 were relatively low, which may indicate less interest in discussing the role of AI in predicting radiation-related side effects. In contrast, Question 41 received the highest score among the three, reflecting interest in AI assistance for detecting underdosed areas which could potentially limit tumour control. RT quality assurance is an essential component of daily practice and has been demonstrated to be critical for optimising disease outcomes and minimising RT-related toxicity [29,30]. Claessens et al. [31] explored the challenges of AI – based RT planning and quality assurance procedures. The authors suggest a process for integrating such systems into everyday practice which includes image registration, auto-segmentation, planning, and quality assurance procedures. Claessens et al. [31] further emphasized that all AI output mandates expert review and corrections if needed (e.g., delineation); they also highlight the challenging role of medical physicists in supervising such systems. Several research groups have been working on models to predict the risk of RT-induced toxicity (e.g., pneumonitis) based on clinical parameters and/or dosimetric parameters. These predictions for normal tissue toxicity are based on different datasets that are available [32–35]. The REQUITE is one of the leading multi-national initiatives for the validation of models and biomarkers predicting risks of late toxicity following RT for different indications [33,36]. It includes breast RT, and the database includes dosimetric parameters, plans, and toxicity outcomes including breast photos to assess cosmesis [36].

Digital/AI is currently utilised in a variety of applications to assist in reducing RT-related toxicity. For instance, manual delineation of specific cardiac substructures is not commonly performed in clinical practice due to the time required and the current focus on broader dose constraints. Recommendations primarily pertain to delineating the whole cardiac structure (e.g., mean heart dose, maximum dose) and, less frequently, the left anterior descending artery (LAD). Cardiac chambers, major arteries, and the three main coronary arteries are rarely delineated, and evidence on their association with RT-related toxicity remains limited. AI-based automated segmentation can aid in delineation particularly as some substructures are not easily defined on RT planning CT scans [37,38]. Causation of later toxicity is challenging to define since it is often multifactorial and necessitates long-term follow-up of a large population. By making such automated segmentation atlases and introducing their routine usage in clinical practice, these supplemental data can serve as a prospectively collected database for future retrospective analysis.

A recent systematic review, Mazo et al. provided an overview of the prediction of breast cancer recurrence using AI techniques. Critically, the authors stated that current publications are significantly lacking in revealing their AI methodologies at different levels such as sampling strategies, data handling, use of feature predictors, and assessment metrics, leading to potential of underdosing a high risk volume [39]. In addition, because the datasets for breast cancer recurrence are scarce, AI model validation and deployment is hindered.

Toolbox recommendations: Breast RT outcomes are not solely focused on disease control. Early breast cancer should be treated curatively, with the preservation of quality of life being a major endpoint for breast cancer survivors. The World Health Organization defined health as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity." Therefore, initiatives such as REQUITE, a multi-national initiatives for the validation of models and biomarkers predicting risks of late toxicity following RT should be supported by the scientific community and by national and international grants to ensure validation and eventual implementation of tools to predict toxicity as an integral component of personalised RT [33,36]. This will allow to treat not only breast cancer but also to reduce treatment related morbidity.

5. Discussion

AI-based systems are increasingly being integrated into various aspects of breast cancer management, including RT optimisation. However, the performance of these systems is highly dependent on the processes underpinning them, necessitating rigorous investment in data refinement to ensure accuracy (Fig. 1). This holds true for all AI applications, and users should maintain a critical approach when evaluating

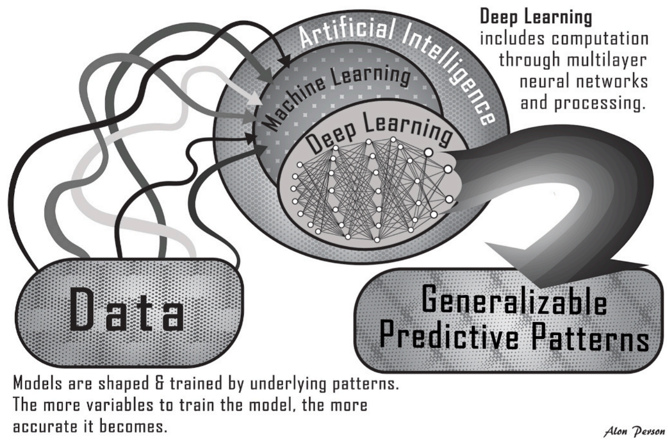

Fig. 1. An illustration of the position of deep learning (DL), comparing with machine learning (ML) and artificial intelligence (AI).

the output for any given use. Furthermore, most publications indicate a lack of transparency regarding how these systems are developed, the information they are based on, the algorithms used, and potential proprietary issues. Such limitations may introduce biases, errors, and raise concerns about data protection [40]. Lack of transparency in AI models or a complex infrastructure of the AI system is seen as one of the main barriers to obtain user trust. In health applications the value of explainable and trustworthy AI has becoming a very important pathway to overcome these difficulties, and lack of trust can be redirected into increasing the staff expertise in identifying potential errors.

AI-based breast cancer RT planning should consider the "human component" (patient) and variations related to patient-specific factors (e.g., anatomical variations, body habitus, and arm mobility for positioning). It should also consider the preferred target volume delineation atlas (e.g., differences between RTOG and ESTRO) and individualization of delineation according to risks of recurrence as recommended by ESTRO guidelines [16].

At present, human expertise cannot be replaced and remains essential at the knowledge and skill level to critically evaluate the output of AI systems. Even so, AI-based planning must account for RT equipment parameters, such as immobilisation devices that may challenge geometric agreement, variations in CT image acquisition and reconstruction, and artefacts such as wires, metal implants, tissue expanders, and machine beam characteristics.

Additionally, the AI planning system, whether implemented in-house or modified from a commercially available provider, should be tailored to the specific devices used in the department (e.g., type of linear accelerator, machine output, respiratory control, surface gating) to address these challenges effectively [20]. All this must ensure compatibility with the users' current and future RT equipment and RT techniques, including MRI-based planning and synthetic CT generation, which may not be compatible with some of the systems currently available [41].

Peer review and quality assurance are engrained in the ethos of RT practice as well as all medical and research practice. It has been demonstrated to reduce treatment-related toxicity, and improve disease outcomes, including in breast cancer RT [42]. The proximity of the breast cancer target volumes to OARs (such as heart and lungs) in patients with good chances at long-term survival merits special attention, as these patients have a life-long risk to develop morbidity from RT. Therefore, breast RT planning necessitates cautious RT planning and delivery.

The use of AI in RT can reduce the workload and improve workflow and may hold even greater value in low- and middle-income countries (LMICs) or in busy radiation departments with insufficiently dimensioned personnel. Serving as a potential solution to the urgent needs to empower the global healthcare and facilitated access to RT [43]. Therefore, projects like the RPA led by specialists from MD Anderson Cancer Center (RPA.mdanderson.org) and developed through partnerships with clinical teams in South Africa, the Philippines, Tanzania, United Kingdom, and the United States, that aim support oncologists in LMICs by reducing the costs associated with RT planning should be reinforced [20]. AI RT planning systems may be of great value in the LMICs where there is a shortage of skilled staff to do RT planning but parallel to that [43], radiation oncology as a community must continue to support improvements in staffing and RT units, including training of the personnel to assure quality care for patients, and not simply rely on "AI" to solve these issues [21,43].

AI is progressively being integrated into various aspects of our lives, including breast cancer management, and patients are progressively putting confidence in AI alongside in their health care providers to improve RT [44,45]. As medical professional users, we must remain vigilant and continually enhance our personal experience and knowledge to remain as the "ground truth" when employing AI, particularly in medical applications such as the treatment planning and delivery of breast cancer patients. Therefore, the basic concepts of the radiation oncology skills should be taught at a high level, including manual delineation of target volumes, treatment planning and peer review. This will allow to develop the AI following an appropriate quality pathway and being applied maintaining a critical mindset. An example for proper development in a critical manner is the development of the DBCG consensus-based auto-segmentation model for lymph node target volumes indicated above [22]. As the work invested in development was meticulous, transparent to the participants, and subjected to revisions and peer review, it allowed to create *trust* in the system abilities by the end user in addition to a high performing system. This eventually will result in reduction of the workload related to breast RT planning, while simultaneously increasing efficiency and productivity. In the future, the EU AI Act could potentially drive more robust evaluations of AI's role in RT, possibly restructuring its development and application. While the EU supervision could lead to a delay in development, it aims to improve patient safety and generate *trust* by assuring high performance of AI-driven solutions [46].

CRediT authorship contribution statement

Orit Kaidar-Person: Writing – original draft, Formal analysis, Writing – review & editing, Investigation, Data curation. **André Pfob:** Supervision, Methodology, Formal analysis, Conceptualization, Writing – review & editing, Project administration, Investigation, Data curation. **Vincenzo Valentini:** Methodology, Conceptualization, Writing – review & editing, Investigation. **Marianne Aznar:** Methodology, Conceptualization, Writing – review & editing, Investigation. **Andre Dekker:** Methodology, Conceptualization, Writing – review & editing, Data curation. **Icro Meattini:** Writing – review & editing, Conceptualization, Methodology. **Jana de Boniface:** Writing – review & editing, Investigation, Methodology, Conceptualization. **David Krug:** Investigation, Writing – review & editing, Conceptualization. **Maria Joao Cardoso:** Writing – review & editing, Conceptualization, Investigation. **Giuseppe Curigliano:** Investigation, Writing – review & editing, Conceptualization. **Peter Dubsky:** Writing – review & editing, Supervision, Project administration, Investigation, Formal analysis, Conceptualization, Validation, Resources, Methodology, Funding acquisition, Data curation. **Philip Poortmans:** Supervision, Investigation, Data curation, Writing – review & editing, Methodology, Formal analysis, Conceptualization.

Acknowledgment

Dr. André Pfob's work is supported by the Digital Oncology Fellowship provided by the German Cancer Research Center (DKFZ) and the National Center for Tumor Diseases (NCT) Heidelberg.

Complete funding and organizational support of the consensus meeting was provided by Hirslanden Klinik St. Anna. The funding source had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Appendix. On behalf of the toolbox consortium[1]

First Name	Family Name	Affiliation
Maggie	Banys-Paluchowski	University Hospital Schleswig-Holstein Campus Lübeck, Germany
Almir	Bitencourt	A.C.Camargo Cancer Center, Brazil
Liesbeth	Boersma	Department of Radiation Oncology (Maastro), GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, The Netherlands
Terri	Coutee	DiepCFoundation.org
Nina	Ditsch	Gynäkologie, Geburtshilfe und Senologie, Medizinische Fakultät, Universität Augsburg, Germany
Maria Alice	BORINELLI-FRANZOI	Cancer Survivorship Group, Inserm Unit 981, Gustave Roussy, Villejuif, France
Dorota	Gabryś	Maria Sklodowska-Curie National Research Institute of Oncology, Gliwice Branch, Poland
Maria Luisa	GASPARRI	1. Faculty of Biomedicine, University of Italian Switzerland (USI), Lugano, Switzerland 2. Dep. Of Gynecology and Obstetrics, Ente Ospedaliero Cantonale (EOC), Lugano, Switzerland
Oreste	Gentilini	1. Vita-Salute San Raffaele University, Milano, Italy 2. IRCCS Ospedale San Raffaele, Milano, Italy
Günther	Gruber	Hirslanden Klinik Zuerich, Switzerland
Nik	Hauser	Brust Zentrum Hirslanden Kliniken Aarau & Cham/Zug, Switzerland
Jörg	Heil	Breast Center Heidelberg, Hospital St. Elisabeth, Heidelberg, Germany
Thorsten	Kuehn	Universitätsfrauenklinik Ulm, Ulm, Germany/Die Filderklinik Filderstadt-Bonlanden, Germany
Vincenzo	L'Imperio	Department of Medicine and Surgery, Pathology, IRCCS Fondazione San Gerardo dei Tintori, University of Milano-Bicocca, Monza, Italy
Ingvil	Mjaaland	Stavanger University Hospital, Stavanger, Norway
Frederique	Penault-Llorca	Department of Pathology, Center Jean Perrin, INSERM U1240, University Clermont Auvergne, Clermont-Ferrand, France
Katja	Pinker	Department of Radiology, Columbia University Irving Medical Center, Vagelos College of Physicians and Surgeons, New York, NY, USA
Mattea	Reinisch	Interdisciplinary Breast Unit, University Hospital Mannheim, Germany
Mann	Ritse	Department of Radiology, Netherlands Cancer Institute (NKI), Amsterdam, The Netherlands; Department of Medical Imaging, Radboud University Medical Center, Nijmegen, The Netherlands
Petra	Steyerova	General university hospital Prague, Czech Republic
Thiemo J. A.	van Nijnatten	Department of Radiology and Nuclear Medicine, GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, the Netherlands
Ines	Vaz	Hirslanden Klinik St. Anna Luzern, Switzerland
Sara E.	Vazquez-Manjarrez	National institute of medical sciences and nutrition Salvador Zubiran, Mexico
Markus	Wallwiener	Halle University Hospital Department for Gynecology, Germany
Walter P	Weber	University of Basel and University Hospital Basel, Switzerland
Isabel T.	Rubio	Breast Cancer Center, Clinica Universidad de Navarra, Madrid, Spain
Duvern	Ramiah	University of the Witwatersrand, Johannesburg, South Africa

References

[1] Kaidar-Person O, Pfob A, Gentilini OD, Borisch B, Bosch A, Cardoso MJ, et al. The Lucerne Toolbox 2 to optimise axillary management for early breast cancer: a multidisciplinary expert consensus. eClinicalMedicine 2023;61:102085.

[2] Dubsky P, Pinker K, Cardoso F, Montagna G, Ritter M, Denkert C, et al. Breast conservation and axillary management after primary systemic therapy in patients with early-stage breast cancer: the Lucerne toolbox. Lancet Oncol 2021;22(1):e18–28.

[3] Dubsky P, Tausch C. Identification of breast cancer patients with pathologic complete response in the breast after neoadjuvant systemic treatment by Pfob et al. Eur J Cancer 2021;143:178–9.

[4] Hoskin PJ. The use of artificial intelligence technologies in cancer care. Clin Oncol 2024.

[5] Robert C, Meyer P, Séroussi B, Leroy T, Bibault JE. Artificial intelligence and radiotherapy: evolution or revolution? Cancer Radiother 2024.

[6] Kochbati L, Zarraa S. Workload patterns in the department of radiotherapy in Salah Azaiez Institute : year 2012. International Atomic Energy Agency (IAEA); 2017. Contract No.: IAEA-CN–250.

[7] Li C, Liu H, Li P, He J, Tian X, Gao W. Applications of artificial intelligence in radiophysics. J Cancer Res Therapeut 2021;17(7):1603–7.

[8] Gibbons E, Hoffmann M, Westhuyzen J, Hodgson A, Chick B, Last A. Clinical evaluation of deep learning and atlas-based auto-segmentation for critical organs at risk in radiation therapy. J Med Radiat Sci 2023;70(Suppl 2):15–25. Suppl 2.

[9] Men K, Zhang T, Chen X, Chen B, Tang Y, Wang S, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. Phys Med 2018;50:13–9.

[10] Schreier J, Attanasi F, Laaksonen H. A full-image deep segmenter for CT images in breast cancer radiotherapy treatment. Front Oncol 2019;9:677.

[11] Liu Z, Liu F, Chen W, Liu X, Hou X, Shen J, et al. Automatic segmentation of clinical target volumes for post-modified radical mastectomy radiotherapy using convolutional neural networks. Front Oncol 2020;10:581347.

[12] Almberg SS, Lervåg C, Frengen J, Eidem M, Abramova TM, Nordstrand CS, et al. Training, validation, and clinical implementation of a deep-learning segmentation model for radiotherapy of loco-regional breast cancer. Radiother Oncol : journal of the European Society for Therapeutic Radiology and Oncology 2022;173:62–8.

[13] Liu Z, Liu F, Chen W, Tao Y, Liu X, Zhang F, et al. Automatic segmentation of clinical target volume and organs-at-risk for breast conservative radiotherapy using a convolutional neural network. Cancer Manag Res 2021;13:8209–17.

[14] Vockner S, Mattke M, Messner IM, Gaisberger C, Zehentmayr F, Ellmauer K, et al. Cone-beam CT segmentation for intraoperative electron radiotherapy based on U-Net variants with transformer and extended LSTM approaches. Cancers (Basel) 2025;17(3).

[15] Choi MS, Choi BS, Chung SY, Kim N, Chun J, Kim YB, et al. Clinical evaluation of atlas- and deep learning-based automatic segmentation of multiple organs and clinical target volumes for breast cancer. Radiother Oncol : journal of the European Society for Therapeutic Radiology and Oncology 2020;153:139–45.

[16] Kaidar-Person O, Offersen BV, Boersma L, Meattini I, Dodwell D, Wyld L, et al. Tricks and tips for target volume definition and delineation in breast cancer: lessons learned from ESTRO breast courses. Radiother Oncol 2021;162:185–94.

[17] Bakx N, Rijkaart D, van der Sangen M, Theuws J, van der Toorn PP, Verrijssen AS, et al. Clinical evaluation of a deep learning segmentation model including manual adjustments afterwards for locally advanced breast cancer. Tech Innov Patient Support Radiat Oncol 2023;26:100211.

[18] van de Sande D, Sharabiani M, Bluemink H, Kneepkens E, Bakx N, Hagelaar E, et al. Artificial intelligence based treatment planning of radiotherapy for locally advanced breast cancer. Phys Imaging Radiat Oncol 2021;20:111–6.

[19] Meixner E, Glogauer B, Kluter S, Wagner F, Neugebauer D, Hoeltgen L, et al. Validation of different automated segmentation models for target volume contouring in postoperative radiotherapy for breast cancer and regional nodal irradiation. Clin Transl Radiat Oncol 2024;49:100855.

[20] Court L, Aggarwal A, Burger H, Cardenas C, Chung C, Douglas R, et al. Addressing the global expertise gap in radiation oncology: the radiation planning assistant. JCO Glob Oncol 2023;9:e2200431.

[21] Court LE, Aggarwal A, Jhingran A, Naidoo K, Netherton T, Olanrewaju A, et al. Artificial intelligence-based radiotherapy contouring and planning to improve global access to cancer care. JCO Glob Oncol 2024;10:e2300376.
[22] Buhl ES, Lorenzen EL, Refsgaard L, Nielsen AWM, Brixen ATL, Maae E, et al. Development and comprehensive evaluation of a national DBCG consensus-based auto-segmentation model for lymph node levels in breast cancer radiotherapy. Radiother Oncol 2024;201:110567.
[23] Yoder T, Hsia AT, Xu Z, Stessin A, Ryu S. Usefulness of EZFluence software for radiotherapy planning of breast cancer treatment. Med Dosim 2019;44(4):339–43.
[24] Kaidar-Person O, Weber WP, Kuhn T, Poortmans P. The forgotten node: axillary surgery mandates expertise. Eur J Surg Oncol 2022;48(9):1922–4.
[25] Coles CE, Haviland JS, Kirby AM, Griffin CL, Sydenham MA, Titley JC, et al. Dose-escalated simultaneous integrated boost radiotherapy in early breast cancer (IMPORT HIGH): a multicentre, phase 3, non-inferiority, open-label, randomised controlled trial. Lancet 2023;401:2124–37. 10394.
[26] Haviland JS, Owen JR, Dewar JA, Agrawal RK, Barrett J, Barrett-Lee PJ, et al. The UK standardisation of breast radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. Lancet Oncol 2013;14(11):1086–94.
[27] Murray Brunt A, Haviland JS, Wheatley DA, Sydenham MA, Alhasso A, Bloomfield DJ, et al. Hypofractionated breast radiotherapy for 1 week versus 3 weeks (FAST-Forward): 5-year efficacy and late normal tissue effects results from a multicentre, non-inferiority, randomised, phase 3 trial. Lancet 2020;395:1613–26. 10237.
[28] Kaidar-Person O, Poortmans P, Salgado R. Genomic-adjusted radiation dose to personalise radiotherapy. Lancet Oncol 2021;22(9):1200–1.
[29] Haas R, Stelmes JJ, Zaffaroni F, Sauvé N, Clementel E, Bar-Deroma R, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of retroperitoneal sarcomas: results from the EORTC 62092-22092 STRASS trial. Cancer 2022;128(14):2796–805.
[30] Kaidar-Person O, Fortpied C, Hol S, Weltens C, Kirkove C, Budach V, et al. The association of internal mammary and medial supraclavicular lymph node radiation technique with clinical outcomes: results from the EORTC 22922/10925 randomised trial. Radiother Oncol 2022;172:99–110.
[31] Claessens M, Oria CS, Brouwer CL, Ziemer BP, Scholey JE, Lin H, et al. Quality assurance for AI-Based applications in radiation therapy. Semin Radiat Oncol 2022; 32(4):421–31.
[32] Massi MC, Gasperoni F, Ieva F, Paganoni AM, Zunino P, Manzoni A, et al. A deep learning approach validates genetic risk factors for late toxicity after prostate cancer radiotherapy in a REQUITE multi-national cohort. Front Oncol 2020;10: 541281.
[33] Seibold P, Webb A, Aguado-Barrera ME, Azria D, Bourgier C, Brengues M, et al. REQUITE: a prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. Radiother Oncol : journal of the European Society for Therapeutic Radiology and Oncology 2019;138:59–67.
[34] West C, Azria D, Chang-Claude J, Davidson S, Lambin P, Rosenstein B, et al. The REQUITE project: validating predictive models and biomarkers of radiotherapy toxicity to reduce side-effects and improve quality of life in cancer survivors. Clin Oncol 2014;26(12):739–42.
[35] Valdes G, Solberg TD, Heskel M, Ungar L, Simone 2nd CB. Using machine learning to predict radiation pneumonitis in patients with stage I non-small cell lung cancer treated with stereotactic body radiation therapy. Phys Med Biol 2016;61(16): 6105–20.
[36] Jaikuna T, Osorio EV, Azria D, Chang-Claude J, De Santis MC, Gutiérrez-Enríquez S, et al. Contouring variation affects estimates of normal tissue complication probability for breast fibrosis after radiotherapy. Breast 2023;72: 103578.
[37] van Velzen SGM, Bruns S, Wolterink JM, Leiner T, Viergever MA, Verkooijen HM, et al. AI-Based quantification of planned radiation therapy dose to cardiac structures and coronary arteries in patients with breast cancer. Int J Radiat Oncol Biol Phys 2022;112(3):611–20.
[38] van Velzen SGM, Gal R, Teske AJ, van der Leij F, van den Bongard D, Viergever MA, et al. AI-Based radiation dose quantification for estimation of heart disease risk in breast cancer survivors after radiation therapy. Int J Radiat Oncol Biol Phys 2022;112(3):621–32.
[39] Mazo C, Aura C, Rahman A, Gallagher WM, Mooney C. Application of artificial intelligence techniques to predict risk of recurrence of breast cancer: a systematic review. J Personalized Med 2022;12(9):1496.
[40] Ferryman K, Mackintosh M, Ghassemi M. Considering biased data as informative artifacts in AI-Assisted health care. N Engl J Med 2023;389(9):833–8.
[41] Palmér E, Karlsson A, Nordström F, Petruson K, Siversson C, Ljungberg M, et al. Synthetic computed tomography data allows for accurate absorbed dose calculations in a magnetic resonance imaging only workflow for head and neck radiotherapy. Phys Imaging Radiat Oncol 2021;17:36–42.
[42] Fisher B, Montague E, Redmond C, Deutsch M, Brown GR, Zauber A, et al. Findings from NSABP protocol no. B-04-comparison of radical mastectomy with alternative treatments for primary breast cancer. I. Radiation compliance and its relation to treatment outcome. Cancer 1980;46(1):1–13.
[43] Zhu H, Chua MLK, Chitapanarux I, Kaidar-Person O, Mwaba C, Alghamdi M, et al. Global radiotherapy demands and corresponding radiotherapy-professional workforce requirements in 2022 and predicted to 2050: a population-based study. Lancet Global Health 2024.
[44] Chan J, Parker L, Carter S, Nickel B, Carroll S. Radiation oncology patients' perceptions of artificial intelligence and machine learning in cancer care: a multi-centre cross-sectional study. Radiother Oncol 2025:110891.
[46] Aboy M, Minssen T, Vayena E. Navigating the EU AI Act: implications for regulated digital medical products. npj Digit Med 2024;7(1):237.
[45] Poortmans PMP, Takanen S, Marta GN, Meattini I, Kaidar-Person O. Winter is over: The use of Artificial Intelligence to individualise radiation therapy for breast cancer. Breast 2020;49:194–200.