Radiotherapy and Oncology 186 (2023) 109748

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

EviGUIDE - a tool for evidence-based decision making in image-guided adaptive brachytherapy for cervical cancer

Stefan Ecker [a,*], Christian Kirisits [a], Maximilian Schmid [a], Johannes Knoth [a], Gerd Heilemann [a], Astrid De Leeuw [b], Alina Sturdza [a], Kathrin Kirchheiner [a], Nina Jensen [c], Remi Nout [d], Ina Jürgenliemk-Schulz [b], Richard Pötter [a], Sofia Spampinato [c], Kari Tanderup [c], Nicole Eder-Nesvacil [a]

[a] Medical University of Vienna, Vienna, Austria; [b] University Medical Centre Utrecht, Department of Radiation Oncology, Utrecht, the Netherlands; [c] Aarhus University Hospital, Department of Oncology, Aarhus, Denmark; [d] Erasmus MC Cancer Institute, University Medical Center Rotterdam, Department of Radiotherapy, Rotterdam, the Netherlands

ARTICLE INFO

Article history:
Received 28 March 2023
Received in revised form 6 June 2023
Accepted 8 June 2023
Available online 16 June 2023

Keywords:
Cervical cancer
IGABT
Decision support
Outcome prediction

ABSTRACT

Purpose: To develop a novel decision-support system for radiation oncology that incorporates clinical, treatment and outcome data, as well as outcome models from a large clinical trial on magnetic resonance image-guided adaptive brachytherapy (MR-IGABT) for locally advanced cervical cancer (LACC).
Methods: A system, called EviGUIDE, was developed that combines dosimetric information from the treatment planning system, patient and treatment characteristics, and established tumor control probability (TCP), and normal tissue complication probability (NTCP) models, to predict clinical outcome of radiotherapy treatment of LACC. Six Cox Proportional Hazards models based on data from 1341 patients of the EMBRACE-I study have been integrated. One TCP model for local tumor control, and five NTCP models for OAR morbidities.
Results: EviGUIDE incorporates TCP-NTCP graphs to help users visualize the clinical impact of different treatment plans and provides feedback on achievable doses based on a large reference population. It enables holistic assessment of the interplay between multiple clinical endpoints and tumour and treatment variables. Retrospective analysis of 45 patients treated with MR-IGABT showed that there exists a sub-cohort of patients (20%) with increased risk factors, that could greatly benefit from the quantitative and visual feedback.
Conclusion: A novel digital concept was developed that can enhance clinical decision- making and facilitate personalized treatment. It serves as a proof of concept for a new generation of decision support systems in radiation oncology, which incorporate outcome models and high-quality reference data, and aids the dissemination of evidence-based knowledge about optimal treatment and serve as a blueprint for other sites in radiation oncology.

The past decades have seen significant advances in the treatment of cervical cancer which is a major cause of mortality for women, especially in low and middle-income countries [1,2]. Driven by excellent clinical results [3,4], combined radiochemotherapy and magnetic resonance (MR) Image-Guided Adaptive Brachytherapy (IGABT) is considered state of the art treatment for patients with locally advanced cervical cancer (LACC) [5,6]. The brachytherapy (BT) component is regarded as a critical element of a successful radiation oncology (RO) treatment, and a growing shift from historic BT practices towards IGABT can be observed [7,8].

In IGABT, and throughout RO, the quality of a treatment plan is commonly assessed using dose-volume histograms (DVH), which summarize the complex volumetric relation between dose distribution, targets, and organs at risk (OAR). It enables the definition of discrete DVH-metrics, at the cost of losing spatial information. A widespread method during the dose optimization process is to compare these metrics to a set of predefined dose-objectives [9], which offers easily interpretable feedback (dose objective achieved/not achieved). Uniquely in RO, for IGABT of LACC there exist recent recommendations that represent international agreement for the prescription, recording and reporting of such dose volume metrics for tumour and OARs [10,11,12,13,14]. A spreadsheet that incorporates this concept along with current dose objectives for multiple endpoints, is distributed alongside report 89 of the international commission on radiation units and measurements (ICRU) and is widely used in clinical practice . It calculates

* Corresponding author at: Medical University of Vienna, Department of Radiation Oncology, Währinger Gürtel 18-20, 1090 Vienna, Austria.
E-mail address: stefan.ecker@meduniwien.ac.at (S. Ecker).

https://doi.org/10.1016/j.radonc.2023.109748
0167-8140/

the total treatment dose from external beam radiotherapy (EBRT) and all (planned, or already irradiated) BT fractions in equieffective dose in 2 Gy per fraction (EQD2) using the linear-quadratic model [10]. The total doses for discrete DVH-parameters for targets and OAR are then graded against their respective dose objectives, which returns a traffic light-style feedback based on their achievement.

Such a tool facilitates interdisciplinary discussion to obtain the best attainable version of a treatment plan for an individual patient. Due to patient diversity, adherence to all dose constraints cannot be assumed a priori and decisions about trade-offs are part of clinical reality. However, the process of this judgement remains challenging because no objective measures to assess the clinical consequences of a dose constraint violation exist. Therefore, current decision-making is driven by expert knowledge of the medical physicist and radiation oncologist. Furthermore, the rigid 'one-size fits all' method of using population based dose objectives is ill-equipped for the future of RO, which is expected to transition towards increasingly personalized medicine [9,15,16,17]. Overcoming these limitations requires prediction models, that estimate the probability of occurrence for clinical events, based on radiation dose and volume, and patient and tumour related features. In RO, dose–response models are categorized into Tumor Control Probability (TCP)-models and Normal Tissue Complication Probability (NTCP)-models. However, assessing the interplay of various clinical endpoints and patient and treatment characteristics is a complex task. Furthermore, integrating outcome prediction models and evidence on optimal treatment into existing clinical workflows demands innovative solutions [18]. In this study, we present a concept for a novel decision-support tool in RO -EviGUIDE-, for the example of MR-IGABT. It utilizes reference data from 1341 patients of the multi-center EMBRACE-I study [3] which recorded a large range of different disease stages, target volumes, prescription concepts and application techniques, therefore representing an exceptional repository of clinical, treatment and outcome data in RO. The interactive tool provides users with personalized predictions for local tumor control and multiple morbidity endpoints and offers feedback on attainable dose distributions for patient subgroups, thereby assisting the treatment planning process.

Methods

Dose-response models

TCP and NTCP models were selected based on recently published results [4,19,20,21,22,23,24], and included clinical relevant endpoints for brachytherapy, with high evidence for dose–effect relationships. Models were constructed based on EMBRACE-I data using uni-and multivariate Cox Proportional Hazards models (Table 1). In total, six different dose–response models were implemented. Endpoints included local tumor control for TCP, and physician-assessed (CTCAE) moderate-to-severe (Grade $\geq$ 2) morbidity events for bladder (pooled bleeding, cystitis, fistula), urinary incontinence, gastrointestinal (pooled proctitis, anal bleeding, rectum bleeding), vaginal stenosis, and flatulence. Predictor variables included the relevant total dose as EQD2 for each endpoint, as well as additional risk factors for local control. See Table D.1 for variable descriptions and used model coefficients. All predictions were calculated at a timepoint of 5 years after treatment. For model evaluation, we report the concordance index (C-index) for time-to-event analysis, using 10-fold cross validation. For each model all results are averaged over 10 random 90/10 splits (training the model with 90% of the EMBRACE-I study population, and hold-out 10% for testing the model).

Patient cohort description

Patients used to build TCP and NTCP models were treated for LACC using combined chemoradiotherapy and BT and were enrolled in the prospective multi-institutional observational cohort study EMBRACE-I (NCT00920920) [3]. Actuarial 5-year local control rates were calculated for multiple patient subgroups (Table D.3) using the Kaplan–Meier estimator and are also referred to in this work as *projected TCP*. To evaluate the utility of the tool, an additional test cohort of 45 patients was analyzed. These patients were treated for LACC at the medical university/general hospital of Vienna between 2016 and 2021, and included all patients treated with high dose rate (HDR) BT and enrolled in the interventional prospective EMBRACE-II study (NCT03617133) [25]. Target definition and dose reporting followed Groupe Européen de Curiethérapie European Society for Radiation Oncology (GEC-ESTRO) recommendations [10,11,12,13,14], and treatment was conducted according to the EMBRACE-II protocol [25] (available at https://www.embracestudy.dk, EK-Nr.2194/2015), which included dose objectives for targets and OAR (Table D.2). To ensure validity of NTCP models, only patients without bladder, rectum or lower/mid vaginal infiltration at diagnosis were included. A summary of dosimetric and patient characteristics is given in Table D.5 and Table 1. The predicted TCP for the clinical treatment plan was compared to the projected TCP of the EMBRACE reference cohort. A subcohort of cases that did not reach the projected TCP levels was selected for additional analysis. The case which is anno-

Table 1
Overview of implemented Cox Proportional Hazards Models showing definition of endpoints, correlated dose, covariates and relevant publications. (G: Grade, OTT: Overall treatment time, A/R: Anus/Rectum).

Endpoint	Dose (EQD2 Gy)	Covariates	C-index[1,2]	Publication
TCP				
Local Tumor Control	$CTV_{HR}D_{90}$	Histopathological Type CTV_{HR} volume OTT Necrosis at diag. (MRI) Corpus uteri infiltration at diag. (MRI)	0.73 (0.09)	Schmid et al. (2023)
NTCP				
Pooled; Bleeding, Cystitis, Fistula G $\geq$ 2	Bladder D_{2cm3}	-	0.59 (0.09)	Spampinato et al., (2021) 312–320
Urinary Incontinence G $\geq$ 2	ICRU-Bladder	-	0.57 (0.11)	Spampinato et al., (2021) 300–308
Pooled; Proctitis, A/R bleeding G $\geq$ 2	Rectum D_{2cm3}	-	0.7 (0.09)	Spampinato et al., (2022)
Vaginal Stenosis G $\geq$ 2	ICRU-RV	-	0.57 (0.06)	Kirchheiner et al., (2016)
Flatulence G $\geq$ 2	Bowel D_{2cm3}	-	0.61 (0.1)	Spampinato et al., (2022)

[1]10-fold CV.
[2]Mean (SD).

tated in Fig. 3 as (1) was excluded from this analysis, as it represents an exceptional case, and is discussed separately.

Investigation of potential TCP improvement

To simulate and assess a potential increase of TCP in the subcohort the clinically used treatment plans were retrospectively improved by uniform scaling of the dose distribution for all BT fractions. For this, the discrete DVH-metrics of the absorbed dose distribution were scaled between 0% and 200%, in discrete steps of 5%, to include a wide range of hypothetical treatment plans. Two approaches to dose escalation were investigated:

1."**With Limits**": maximize TCP w.r.t. EMBRACE-II dose objectives (hard constraints, Table D.2).

2."**Without Limits**": escalate dose until projected TCP is achieved.

For the first algorithm, the target dose CTV_{HR} D_{90} was additionally limited to 100 Gy $EQD2_{10}$ to ensure clinically feasible plans. Differences between the clinical and improved treatment plans were reported in terms of TCP or NTCP and dose in EQD2 for all endpoints. Based on the scale factor that produced the best plan, patients were categorized into two categories:" No Change" (scale factor = 100%), and"Dose Escalation" (scale factor > 100%). Analysis was performed in R version 4.0.2 [26], using a custom script that first calculated the scaled versions of each treatment plan, and then identified the *best* plan according to the objective functions above.

Results

A high-level overview of the tool architecture can be found in Fig. 1. Fig. 2 shows screenshots of the graphical user interface (GUI). It consists of five major modules, that are able to process treatment plans from patients treated with HDR BT. Information about the two technical modules, *Data Import and Processing* and *EQD2 Dose Documentation* is given in Appendix B.

The first module - Tumour and Treatment Characteristics - allows for user-defined selection of multiple tumour, and treatment characteristics. The selection of parameters influences either predictions for local tumor control and/or the selection of similar patients for the dosimetric reference cohort. An overview of all available parameters and their respective influence can be found in Table D.7.

The next module - TCP-NTCP assessment and simulation - calculates and visualizes the personalized risk for local failure and side-effects, based on the loaded treatment plan and selected characteristics. Model evaluation based on 10-fold cross validation using the EMBRACE-I dataset resulted in C-indices ranging from 0.57 (urinary incontinence, vaginal stenosis) to 0.73 (local tumor control). The aim of this module is to assist the decision-making process of finding the plan with best possible balance between risks of local failure and harm to OAR for a given patient. The interplay between doses and predicted probabilities can be visualized in two different ways. First, traditional dose–response curves are displayed, which show the dependence of each endpoint (y-axis) on their respective dose in EQD2 (x-axis). Predicted values are indicated with a black line, and their 95% confidence intervals are represented in gray. Red and orange dashed vertical lines represent relevant dose objectives from the EMBRACE-II protocol (Table D.2). Horizontal and vertical green lines highlight the loaded treatment plan. In total, users can visualize up to six different dose–response curves simultaneously (Fig. 2D). Second, the treatment plan can be assessed in the space of NTCP and TCP as $\times$ and y-axes, respectively. Five versions of this plot, one for each morbidity endpoint, can be visualized. Fig. 3 illustrates the structure of this plot for bladder G $\geq$ 2 events for the full test cohort.

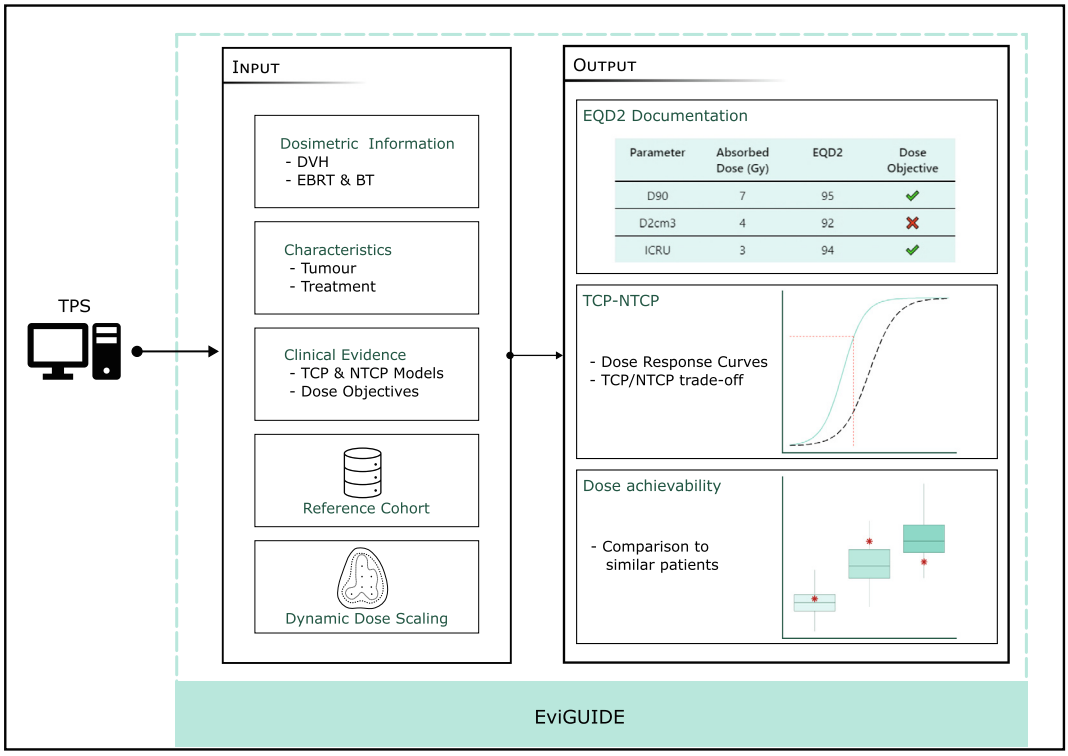

Fig. 1. Schematic overview of the EviGUIDE software. A link to a commercial TPS provides dosimetric information about the case. Tumour and treatment characteristics can be selected by the user via a graphical user interface. Three different outputs are produced; a tabular summary of relevant EQD2 parameters, TCP and NTCP curves that allow for assessment of radiobiological trade-off, and summary statistics that compare the current case to a sub-cohort of similar patients. Input controls allow uniform scaling of the dose distribution, to dynamically create and compare derivative treatment plan options.

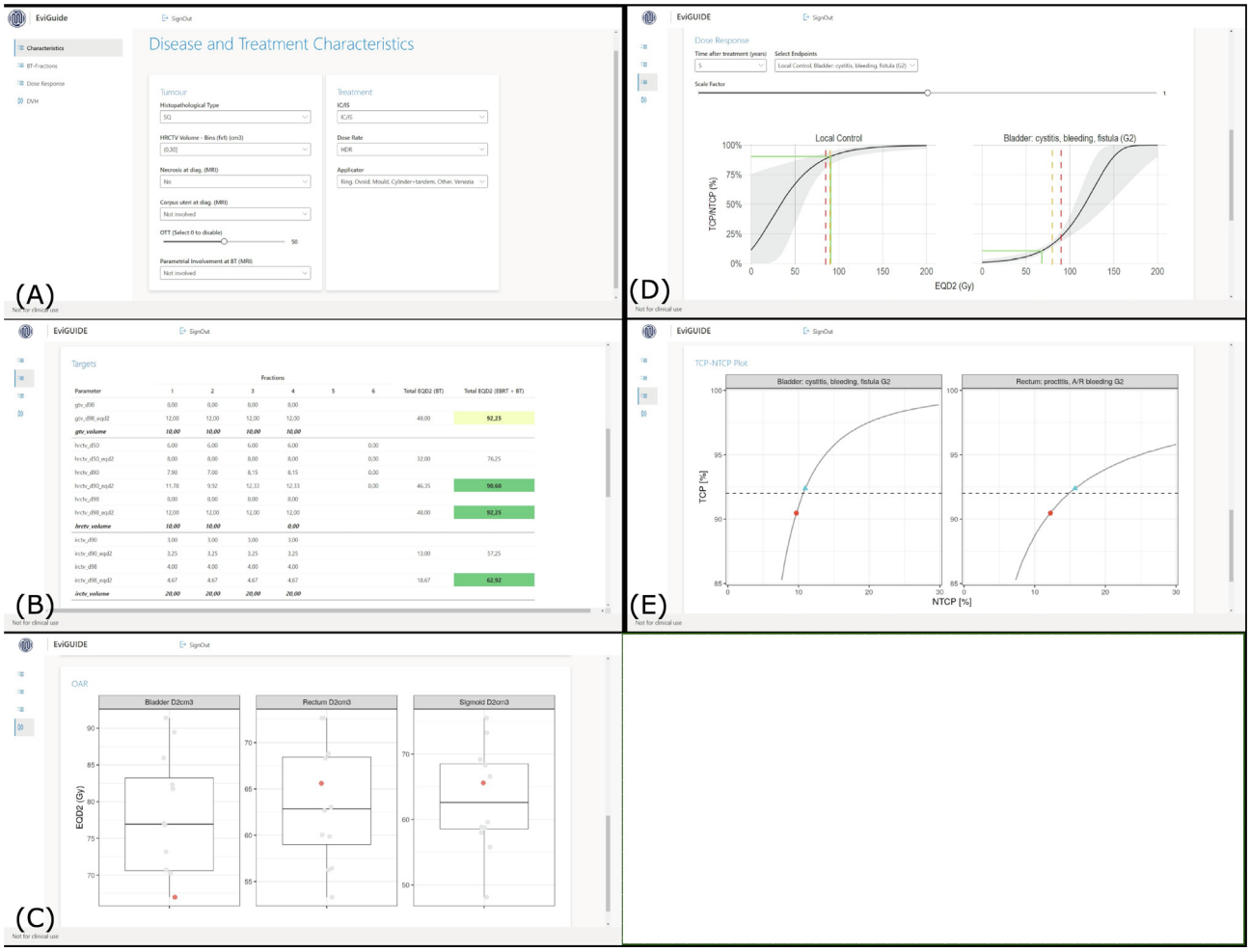

Fig. 2. Screenshots of EviGUIDE graphical user interface, showing the different modules. (A) User interface to select tumour and treatment characteristics. (B) EQD2 dose documentation table, color-coded according to dose objectives. (C) DVH statistics - boxplots summarizing target and OAR doses and comparing the current patient (red dot) with a reference cohort of similar patients (gray). (D) Top: Input controls for selecting a discrete scale factor, time after treatment, and endpoints. Bottom: Dose-response curves of two selected endpoints. Green line: loaded treatment plan; dashed lines: dose objectives (Tab. D.2) (E) TCP-NTCP plots for two morbidity endpoints. Red points represent the loaded treatment plan. Blue triangles a scaled, hypothetical version of the treatment plan. Horizontal dashed line represents projected TCP for this case.

The location of a treatment plan in the TCP-NTCP space, for a specific patient is marked by a point. By uniformly scaling the absorbed dose of all BT fractions between 0% and 200%, characteristic curves are generated. These curves represent all theoretically attainable points of therapeutic outcome, with the given treatment plan. The shape of the curve is determined by (i) the underlying TCP and NTCP models, (ii) loaded treatment plan and (iii) patient characteristics. Evidently, the most favorable configurations are represented by points in the upper left quadrant of the NTCP-TCP space (high local control and limited side-effects). The curve monotonically increases with escalating dose. Based on data from EMBRACE-I it is possible to determine what TCP could be realistically expected for different patient subgroups. Table D.3 shows Kaplan-Meier estimates for local control, i.e., the *projected TCP*. If the treatment plan meets or exceeds the projected TCP, its representing point in the NTCP-TCP plot is highlighted in blue. Otherwise, it is highlighted in red. Users are able to interactively select discrete scale factors, and directly compare the difference between the original and scaled treatment plan, in terms of total dose in EQD2 and outcome. The results are summarized in a data table. Both visualization options dynamically change according to the manually selected scale factor and provide visual cues to compare the treatment plan options.

The final module - DVH Statistics – puts the plan and patient in relation to a large reference cohort. For the loaded treatment plan, a reference cohort of similar patients can be defined by selecting relevant patient and treatment characteristics (Table D.7). The sub-cohort of similar patients is constructed by filtering the overall cohort by these parameters. Boxplots of BT target (CTV_{HR} D_{90}, CTV_{IR} D_{98}, GTV_{res} D_{98}) and OAR (Bladder D_{2cm3}, Rectum D_{2cm3}, Sigmoid D_{2cm3},Bowel D_{2cm3}, ICRU recto-vaginal point) doses, as well as target volumes (CTV_{HR}, CTV_{IR}, GTV_{res}) are generated. Plots are dynamically adjusted to the reference cohort, which is shown in gray. The treatment plan under investigation is highlighted via a red datapoint. Boxplots show the median, 25th and 75th quartiles, and either the test cohort of 45 patients or the entire EMBRACE study population could be selected as a reference cohort.

Overall, the majority of patients treated with state-of-the-art MR-IGABT (33/45, 73%) achieved their projected TCP using the clinical treatment plans. For the subcohort of 11 patients that did not reach this mark potential for plan adaptation was further investigated using the uniform dose scaling approach. TCP-maximization by dose scaling "With Limits" indicated that for 9/11 (82%), an average increase of TCP from 91% to 93% could be achieved, requiring average CTVHR D90 dose escalation from 91 Gy to 98 Gy $EQD2_{10}$. This benefit for local tumor control would

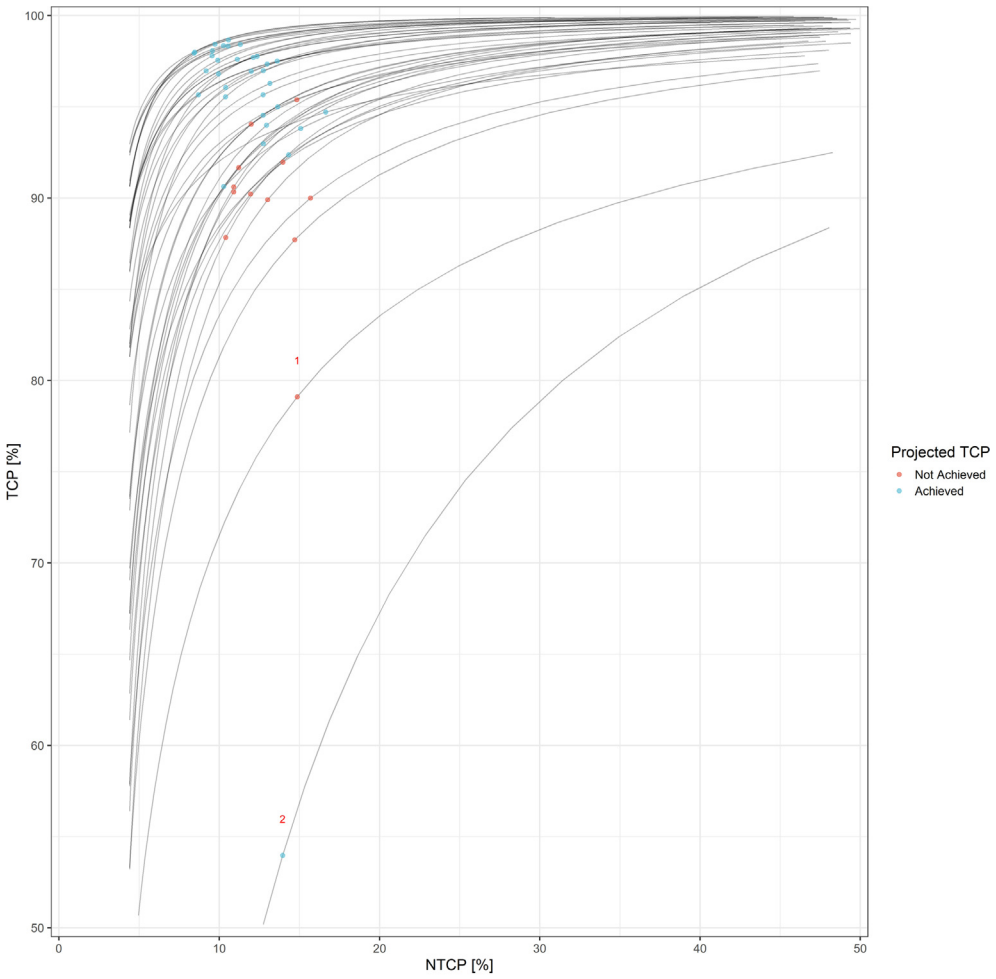

Fig. 3. Plot of TCP and bladder G $\geq$ 2 NTCP, for all patients in the test cohort. Each point represents a clinical treatment plan. Predictions were based on dosimetric information, patient characteristics and Cox Proportional Hazards Models, 5-years after treatment (Table 1). Curves represent all theoretically attainable points of therapeutic outcome, by uniform scaling of the dose distribution between 0% and 200%. Plans highlighted in blue meet or exceed their projected TCP based on outcome data from EMBRACE-I. Otherwise plans are marked in red. Two outlier patients were annotated as (1) and (2) for in-depth discussion.

come at the cost of additional NTCP and dose to OAR of < 1% – 2%, and 1–5 Gy $EQD2_3$, respectively. Table 2 lists detailed statistics for all TCP and NTCP endpoints. For two patients no improved version of the clinical treatment plan was found using this method. Simulated TCP-maximization by dose scaling "Without Limits" would on average increase the predicted TCP from 91% to 94%. According to the simulation, the average CTV_{HR} D_{90} would have to be escalated from 92 Gy to 103 Gy $EQD2_{10}$, and cost of additional NTCP and dose to OAR of 1% – 4%, and 2–9 Gy $EQD2_3$, respectively (see Table D.4). Fig. 4 shows the NTCP-TCP graphs for patients of the sub-cohort. It compares the clinically used treatment plans, with their derived versions from the algorithm "With Limits" and "Without Limits". Each row shows one of the five NTCP endpoints, illustrating the interplay of TCP and NTCP as a function of prescribed dose.

Discussion

While concepts to include NTCP models into RO for sites like head and neck cancer have been proposed [27,28], limited information is available about dedicated prediction models for cervical cancer, and standard prognostic tools are generally used retrospectively and implemented using nomograms [29,30,31]. Nevertheless, the emergence of artificial intelligence (AI) to improve patient care has significantly influenced healthcare research in the past decade [32,33,34,35], and access to large multi-national datasets and more powerful models has increased interest in using AI for outcome prediction or other tasks in RO [36,37,38,39]. However, this development also presents new challenges. It has been highlighted that model interpretability, and how humans will interact with these algorithms in clinical practice, are pressing questions in RO and healthcare in general [33,40,41,42]. Intuitive and interpretable interfaces, such as dashboards, are needed to communicate the results of these models in a safe way, and ensure clear communication of their limitations [18,43,44]. On the other hand, such tools would also open up new possibilities for shared decision-making between patients and clinicians. Personal preferences of the patient, which may be influenced by age, expected outcome or future desires, should always be taken into account by the treatment team and may not always reflect dosimetric objectives [45,46]. The tool that was developed in this study can serve as a proof of concept for such a new generation of decision support systems in RO.

Furthermore, it is important to emphasize that for MR-IGABT, treatment plan *optimization* is not a process of minimizing or maximizing a specific metric. In clinical reality the determination of the *optimal* treatment plan is complex. Target and OAR prioritization, evidence on dose–effect relationships, implant quality and practical feasibility of achievable dose [47], should all be considered. In this regard, EviGUIDE can provide valuable support for going

Table 2
Summary statistics for patients that did not achieve their projected TCP based on outcome data from EMBRACE-I. The difference between clinical, and improved treatment plans "With Limits" (i.e. maximize TCP w.r.t. EMBRACE-II dose objectives) is described. For each endpoint, average (N)TCP, and dose in EQD2 Gy including standard deviation is reported. For 2 patients no better version of a treatment plan could be found. Statistical significance $p < 0.05$.

Endpoint	**Dose Escalation**					**No Change**	
	clinical, N = 9[1]	**improved**, N = 9[1]	**Difference**[2]	**95% CI**[2,3]	**p-value**[2]	**clinical**, N = 2[1]	**improved**, N = 2[1]
Local Control							
TCP (%)	91 (2)	93 (2)	2	2, 1	< 0.001	89 (1)	89 (1)
HRCTV D90 (EQD2 Gy)	91 (2)	98 (2)	7	8, 5	< 0.001	92 (6)	92 (6)
Pooled; Bleeding, Cystitis, Fistula G $\geq$ 2							
NTCP (%)	13 (2)	15 (2)	2	3, 2	< 0.001	12 (2)	12 (2)
Bladder D2cc (EQD2 Gy)	77 (5)	82 (5)	5	7, 4	< 0.001	74 (5)	74 (5)
Urinary Incontinence G $\geq$ 2							
NTCP (%)	14 (2)	14 (2)	1	1, 0	< 0.001	16 (5)	16 (5)
ICRU-Bladder (EQD2 Gy)	68 (11)	72 (13)	4	5, 2	< 0.001	79 (25)	79 (25)
Pooled; Proctitis, A/R bleeding G $\geq$ 2							
NTCP (%)	9 (4)	11 (6)	2	4, 1	0.015	9 (4)	9 (4)
Rectum D2cc (EQD2 Gy)	61 (7)	64 (8)	3	4, 1	0.002	61 (7)	61 (7)
Vaginal Stenosis G $\geq$ 2							
NTCP (%)	23 (4)	24 (4)	1	1, 0	0.004	24 (2)	24 (2)
ICRU-RV (EQD2 Gy)	55 (7)	57 (8)	2	2, 1	0.001	56 (4)	56 (4)
Flatulence G $\geq$ 2							
NTCP (%)	8 (1)	8 (1)	0	0, 0	0.014	9 (2)	9 (2)
Bowel D2cc (EQD2 Gy)	50 (5)	51 (6)	1	1, 0	0.007	53 (8)	53 (8)

[1]Mean (SD).
[2]Paired t-test.
[3]CI = Confidence Interval.

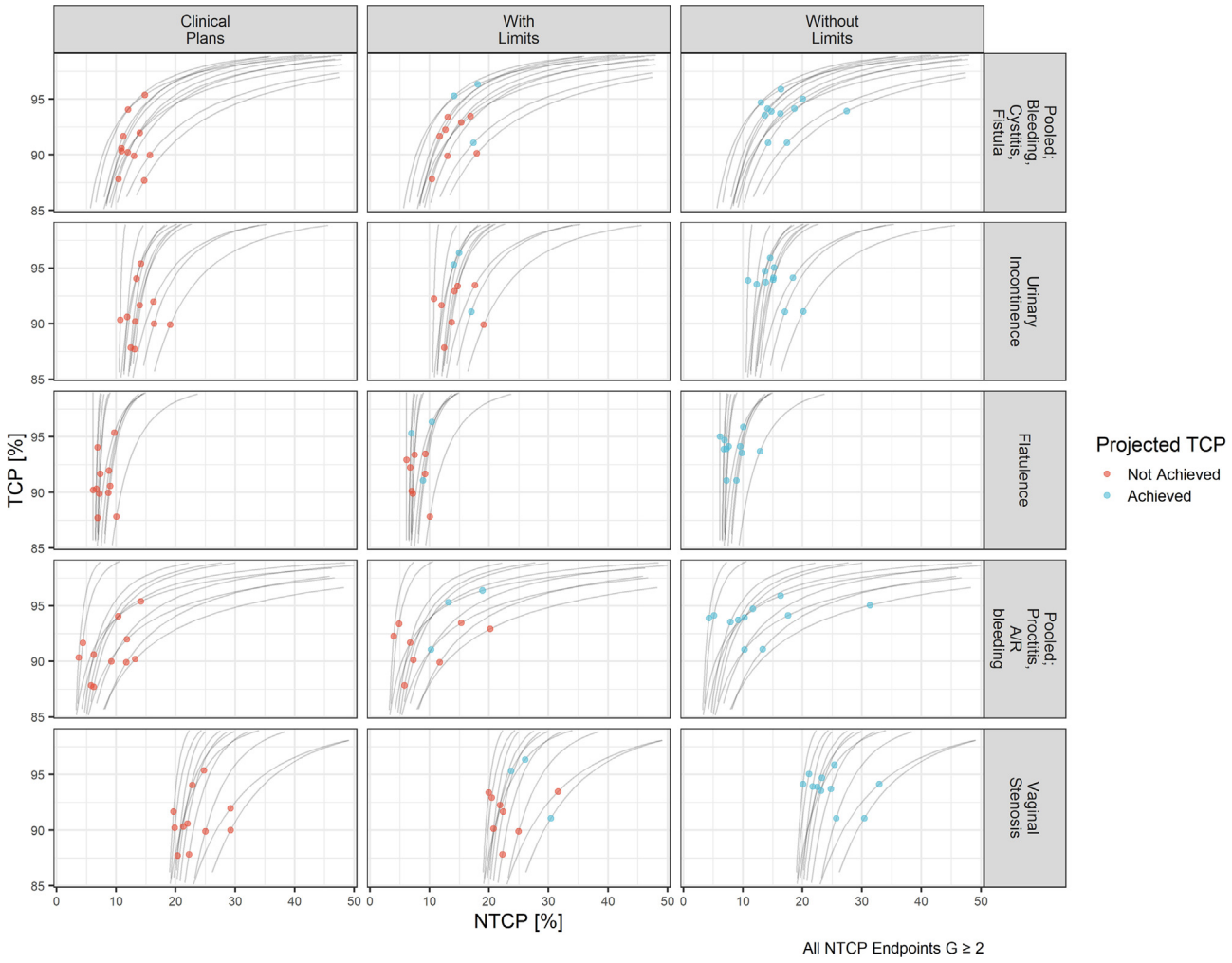

Fig. 4. TCP-NTCP plots showing patients that did not achieve their projected TCP using the clinical treatment plan (left). Hypothetical plans obtained by algorithm "With Limits" (middle) and "Without Limits" (right) are shown for comparison. Each row represents a different morbidity endpoint with Grade $\geq$ 2.

beyond expert knowledge based on individual clinical experience only.

Using evidence-based dose response curves, as well as the TCP-NTCP space provides advanced insights about the clinical impact of a treatment plan. By visualizing multiple TCP-NTCP trade-offs, users gain valuable information about the cost-benefit of possible treatment options. Interactive scaling of the BT dose distribution allows users to quickly explore different possibilities of an existing treatment plan. Appendix Figure E.1 illustrates the tool concept based on a clinical example case.

The concept of NTCP-TCP graphs was already proposed some decades ago [48,49], and could also be used to define objective metrics for an ideal operating point of a treatment plan [50]. In addition, the GUI enables users to easily select relevant patient and treatment features, allowing for individualization of TCP, and selection of a reference cohort. The summarizing boxplots provide an interpretable overview of how the current treatment plan relates to other similar patients, and aims to provide feedback on realistically achievable doses, based on previous plans. In this version of EviGUIDE the selection of similar patients by filtering of selected features is rather simplistic. It cannot take into account patient anatomy from imaging data i.e., geometrical features, as this modality has not been collected during the EMBRACE studies, however the large pool of over 1000 available reference cases still provides valuable feedback. In future versions, more advanced data-driven methods to improve patient clustering would be of interest to improve selection of similar patients [32]. Furthermore, EviGUIDE not only allows the assessment of the current plan but can assist in the planning of subsequent BT fractions. Based on the feedback from the TCP/NTCP curves, and comparison with a reference cohort of similar patients, practitioners can assess whether a change of practice for the remaining treatment is warranted. The comparison with reference data from EMBRACE-I could be expanded with additional parameters, such as the average number of active needles that were used in cases with similar disease extent, which could help guide the need of, or addition of further, interstitial needles in subsequent fractions.

In practice, after an initial treatment plan was created, the dosimetric data can be sent to EviGUIDE via the TPS-specific software link. The medical physicist and radiation oncologist can use the interactive tool to assess the quality of the treatment plan based on large clinical evidence and consider possible adaptations to improve the projected treatment outcome. After decision on the final treatment plan, the tool can be used to create a filled-out spreadsheet (as published in ICRU 89) to document the treatment. This also eliminates the need for manual dose documentation and provides a time-saving opportunity while reducing the risk of typing errors. In the current version, information flow was a one-way-street from TPS to EviGUIDE. Nevertheless, for future applications, the information about TCP and NTCP could also be returned to the TPS or other software to assist automatic planning algorithms [51], thus closing the dose optimization loop. One of the limitations of the current software prototype is that no full BT dose plan optimization, i.e., modifying the 3D dose distribution by adaptation of the individual dwell time distribution, was triggered by the feedback from EviGUIDE. However, the simple simulation of potential TCP and NTCP modification by scaling of the BT dose distribution serves as a surrogate for future sophisticated techniques for informed automated plan optimization. One of the advantages of the current tool is the independence from a commercial TPS, as interfaces to all available TPS for semi-automated DVH import can be easily established.

Results from the retrospective evaluation of the test cohort show that overall patients could expect favorable treatment outcome, in line with benchmarks from EMBRACE-I. This is an expected result, as these patients were treated according to the EMBRACE-II protocol. However, even among these patients that were treated with a high standard of MR-IGABT, there exists a sub-cohort of cases where the planning team could have benefited from the quantitative and visual feedback provided by EviGUIDE. Improving TCP of treatment plans within existing EMBRACE-II dose constraints revealed that for 20% (9/45) of patients in the test cohort, there could have been potential to exhaust dose limits. On average this would have resulted in a 2% increased TCP, at the cost of increasing Grade $\geq$ 2 side-effect probability for each organ in the range of < 1–2%, on average (Table 2). However, as is revealed in Fig. 4, only three of these patients would have been able to achieve their projected TCP using this method. Further dose escalation beyond existing dose limits demonstrated an average gain in TCP of 3%, while increasing the risk for Grade $\geq$ 2 side-effects for each organ in the range of 1–4%, on average (Table D.4). This indicates that while not all patients would be able to meet their expected TCP in clinical reality, as they were constrained by their dose to OAR, especially bladder, rectum and vagina (Fig. 4), some cases showed potential for individualized dose escalation. In addition, the range for optimization would likely be higher for inexperienced centers that are in the process of adapting MR-IGABT. While it is acknowledged that the uniform scaling of discrete DVH-metrics between 0% and 200% may generate some hypothetical treatment plans that fall outside clinically relevant dose ranges, these plans still contribute valuable insights to interpreting the characteristic TCP/NTCP curves (Fig. 3). By including these data points, the curve's shape and steepness become more pronounced, serving as indicators of how closely a current treatment plan approaches the limit for a specific OAR.

As it is illustrated in Fig. 3, the distribution of local tumor control probabilities varies among patients. Fig. 4 further shows the complex interplay between different clinical endpoints and highlights the benefits of visualizing all possible combinations side-by-side. This feature of EviGUIDE enables users to focus on patient-specific TCP-NTCP trade-offs, and to assess whether a change in dose prescription would be justified despite non-adherence to current dose objectives. This was especially relevant in the two prominent outliers which were annotated as (1) and (2) in Fig. 3. The two cases are described in detail in Appendix C, and also highlight existing limitations of the current version of EviGUIDE. First, although the TCP and NTCP models represent the state-of-the-art, significant uncertainties remain, particularly for small patient subgroups and in dose ranges beyond those typically used in clinical practice (CTV_{HR} D_{90} 85–95 Gy $EQD2_{10}$). In addition, discrete DVH-parameters are used as a surrogate for the 3D dose distribution which limits predictive accuracy. These factors are reflected in the relatively low C-indices of some morbidity endpoints (Table 1). Finally, holistic assessment of LACC treatment requires consideration of other patient and disease endpoints beyond local tumor control, such as baseline morbidity or systemic disease, which should be incorporated once data and models become available. It is therefore emphasized that this is a proof of concept, and translation into clinical practice would require prior external validation and a prospective clinical study to demonstrate its ability to improve medical decision-making when compared to state-of-the-art clinical judgement alone.

To conclude, this study presents a novel digital concept (EviGUIDE) to exploit the wealth of clinical, treatment and outcome data, as well as sophisticated TCP and NTCP models from a large clinical trial on MRI-IGABT for cervical cancer based on internationally agreed upon metrics, both for tumour, target and OAR assessment. EviGUIDE represents therefore an evidence-based decision support system for personalized dose and volume prescription. It facilitates access to this complex data resource and complements existing treatment planning workflows by accelerating the radiotherapy workflow, delivering individualized and inter-

pretable predictions on expected outcome and providing feedback on achievable dose distributions. It will serve as a blueprint for other sites in radiation oncology. It has to be emphasized that the application of an evidence-based decision information tool as proposed in this study should be based on TCP-NTCP models derived from patient, tumor and treatment data of high quality for multiple endpoints and covering a wide range of possible treatments. To our knowledge the EMBRACE studies, which provide such a high-quality database, enable for the first time the implementation of such a concept in one comprehensive application.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project was supported by the Austrian Science Fund, Austria (FWF, project number KLI695-B33).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2023.109748.

References

[1] Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. Lancet 2019;393:169–82. https://linkinghub.elsevier.com/retrieve/pii/S014067361832470X.
[2] Rodin D et al. Scale-up of radiotherapy for cervical cancer in the era of human papillomavirus vaccination in low-income and middle-income countries: a model-based analysis of need and economic impact. Lancet Oncol 2019;20:915–23. https://linkinghub.elsevier. com/retrieve/pii/S1470204519303080.
[3] Pötter R et al. MRI-guided adaptive brachytherapy in locally advanced cervical cancer (EMBRACE-I): a multicentre prospective cohort study. Lancet Oncol 2021;22:538–47. https://linkinghub.elsevier.com/retrieve/pii/S1470204520307531.
[4] Schmid, M. P. *et al.* Risk Factors for Local Failure Following Chemoradiation and Magnetic Resonance Image–Guided Brachytherapy in Locally Advanced Cervical Cancer: Results From the EMBRACE-I Study. *Journal of Clinical Oncology* JCO.22.01096 (2023). URL https: //ascopubs.org/doi/10.1200/JCO.22.01096.
[5] Sturdza AE, Knoth J. Image-guided brachytherapy in cervical cancer including fractionation. Int J Gynecol Cancer 2022;32:273–80. https://doi.org/10.1136/ijgc-2021-003056.
[6] Cibula D et al. The European Society of Gynaecological Oncology/European Society for Radiotherapy and Oncology/European Society of Pathology Guidelines for the Management of Patients with Cervical Cancer. Virchows Arch 2018;472:919–36. https://doi.org/10.1007/s00428-018-2362-9.
[7] Han K, Milosevic M, Fyles A, Pintilie M, Viswanathan AN. Trends in the Utilization of Brachytherapy in Cervical Cancer in the United States. Int J Radiat Oncol Biol Phys 2013;87:111–9. , https://linkinghub.elsevier. com/retrieve/pii/S0360301613005956.
[8] Tan L-T et al. Education and training for image-guided adaptive brachytherapy for cervix cancer—The (GEC)-ESTRO/EMBRACE perspective. Brachytherapy 2020;19:827–36. https://linkinghub.elsevier.com/retrieve/pii/S1538472120301306.
[9] Hernandez V et al. What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans. Radiother Oncol 2020;153:26–33. https://linkinghub.elsevier.com/retrieve/pii/S0167814020308136.
[10] ICRU report 89, Prescribing, recording, and reporting brachytherapy for cancer of the cervix. *J ICRU*. 2013; **13** (NP)
[11] Haie-Meder C et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (I): concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV. Radiother Oncol 2005;74:235–45. https://linkinghub.elsevier.com/retrieve/pii/S0167814004005791.
[12] Pötter R et al. Recommendations from gynaecological (GYN) GEC ESTRO working group (II): Concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy—3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. Radiother Oncol 2006;78:67–77. https://linkinghub.elsevier.com/retrieve/ pii/S0167814005005463.
[13] Hellebust TP et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group: Considerations and pitfalls in commissioning and applicator reconstruction in 3D image-based treatment planning of cervix cancer brachytherapy. Radiother Oncol 2010;96:153–60. https://linkinghub.elsevier.com/retrieve/pii/S0167814010003683.
[14] Dimopoulos JC et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (IV): Basic principles and parameters for MR imaging within the frame of image based adaptive cervix cancer brachytherapy. Radiother Oncol 2012;103:113–22. https://linkinghub.elsevier.com/retrieve/pii/S016781401100764X.
[15] Fiorino C et al. Grand challenges for medical physics in radiation oncology. Radiother Oncol 2020;153:7–14. https://linkinghub.elsevier.com/retrieve/pii/S0167814020308379.
[16] Liauw SL, Connell PP, Weichselbaum RR. New paradigms and future challenges in radiation oncology: An update of biological targets and technology. Sci Transl Med 2013;5. https://www.science.org/doi/10.1126/scitranslmed.3005148.
[17] Bibault J-E, Giraud P, Burgun A. Big Data and machine learning in radiation oncology: State of the art and future prospects. Cancer Lett 2016;382:110–7.
[18] Overgaard J et al. Personalised radiation therapy taking both the tumour and patient into consideration. Radiother Oncol 2022;166:A1–5. https://linkinghub. elsevier.com/retrieve/pii/S0167814022000147.
[19] Spampinato S et al. Risk factors and dose-effects for bladder fistula, bleeding and cystitis after radiotherapy with imaged-guided adaptive brachytherapy for cervical cancer: An EMBRACE analysis. Radiother Oncol 2021;158:312–20. , https://linkinghub.elsevier. com/retrieve/pii/S0167814021000207.
[20] Spampinato S et al. Importance of the ICRU bladder point dose on incidence and persistence of urinary frequency and incontinence in locally advanced cervical cancer: An EMBRACE analysis. Radiother Oncol 2021;158:300–8. , https://linkinghub.elsevier.com/ retrieve/pii/S0167814020308392.
[21] Spampinato, S. *et al.* Severity and Persistency of Late Gastrointestinal Morbidity in Locally Advanced Cervical Cancer: Lessons Learned From EMBRACE-I and Implications for the Future. *International Journal of Radiation Oncology*Biology*Physics* **112**, 681–693 (2022). URL https: //linkinghub.elsevier.com/retrieve/pii/S0360301621029151.
[22] Westerveld H et al. Dose-effect relationship between vaginal dose points and vaginal stenosis in cervical cancer: An EMBRACE-I sub-study. Radiother Oncol 2022;168:8–15. https://linkinghub.elsevier.com/retrieve/pii/S0167814021090824.
[23] Mazeron R et al. Dose–volume effect relationships for late rectal morbidity in patients treated with chemoradiation and MRI-guided adaptive brachytherapy for locally advanced cervical cancer: Results from the prospective multicenter EMBRACE study. Radiother Oncol 2016;120:412–9.
[24] Kirchheiner K et al. Dose–effect relationship and risk factors for vaginal stenosis after definitive radio(chemo)therapy with image-guided brachytherapy for locally advanced cervical cancer in the EMBRACE study. Radiother Oncol 2016;118:160–6. , https://linkinghub. elsevier.com/retrieve/pii/S0167814016000025.
[25] Pötter R et al. The EMBRACE II study: The outcome and prospect of two decades of evolution within the GEC-ESTRO GYN working group and the EMBRACE studies. Clin Transl Radiat Oncol 2018;9:48–60. , https://linkinghub.elsevier.com/ retrieve/pii/S2405630817300757.
[26] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2021). URL https://www.R-project.org/.
[27] Langendijk JA et al. Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach. Radiother Oncol 2013;107:267–73. https://linkinghub.elsevier.com/retrieve/pii/S0167814013002193.
[28] Van den Bosch L et al. Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: A new concept for individually optimised treatment. Radiother Oncol 2021;157:147–54. https://linkinghub.elsevier.com/retrieve/pii/S0167814021060217.
[29] Sturdza AE et al. Nomogram predicting overall survival in patients with locally advanced cervical cancer treated with radiochemotherapy including image-guided brachytherapy: A retro-EMBRACE study. Int J Radiat Oncol Biol Phys 2021;111:168–77. https://linkinghub.elsevier.com/retrieve/pii/S0360301621004041.
[30] Raymond, E. *et al.* An appraisal of analytical tools used in predicting clinical outcomes following radiation therapy treatment of men with prostate cancer: a systematic review. *Radiation Oncology* **12**, 56 (2017). URL http://ro-journal.biomedcentral.com/articles/10. 1186/s13014-017-0786-z.
[31] He B et al. Prediction models for prognosis of cervical cancer: systematic review and critical appraisal. Front Public Health 2021;9:. https://doi.org/10.3389/fpubh.2021.654454/full654454.
[32] Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. Med Phys 2019;46:2760–75. https://doi.org/10.1002/mp.13526.
[33] Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022;28:31–8. , https://www.nature.com/articles/s41591-021-01614-0.
[34] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018;18:500–10. , http://www.nature.com/articles/ s41568-018-0016-5.

[35] Sahiner B et al. Deep learning in medical imaging and radiation therapy. Med Phys 2019;46:e1–e36. https://onlinelibrary.wiley.com/doi/10.1002/mp.13264.
[36] Luo Y, Chen S, Valdes G. Machine learning for radiation outcome modeling and prediction. Med Phys 2020;47. , https://onlinelibrary.wiley.com/doi/10.1002/mp.13570.
[37] Appelt A, Elhaminia B, Gooya A, Gilbert A, Nix M. Deep learning for radiotherapy outcome prediction using dose data – A review. Clin Oncol 2022;34:e87–96. https://linkinghub.elsevier.com/retrieve/pii/S0936655521004623.
[38] Deist TM et al. Machine learning algorithms for outcome prediction in (chemo) radiotherapy: An empirical comparison of classifiers. Med Phys 2018;45:3449–59. https://doi.org/10.1002/mp.12967.
[39] Lambin P et al. Predicting outcomes in radiation oncology—multifactorial decision support systems. Nat Rev Clin Oncol 2013;10:27–40. http://www.nature.com/ articles/nrclinonc.2012.196.
[40] MI in Healthcare Workshop Working Group *et al.* Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency. *npj Digital Medicine* **3**, 47 (2020). URL http://www.nature.com/articles/s41746-020-0254-2.
[41] Marwaha JS, Landman AB, Brat GA, Dunn T, Gordon WJ. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. npj Digital Med 2022;5:13. https://www.nature.com/articles/s41746-022-00557-1.
[42] Walsh S et al. Decision support systems in oncology. JCO Clin Cancer Informatics 2019;1–9. https://doi.org/10.1200/CCI.18.00001.
[43] Banegas-Luna AJ et al. Towards the interpretability of machine learning predictions for medical applications targeting personalised therapies: A cancer case survey. Int J Mol Sci 2021;22:4394. , https://www.mdpi.com/1422-0067/22/9/4394.
[44] Morin O et al. An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication. Nat Cancer 2021;2:709–22. http://www.nature.com/articles/s43018-021-00236-2.
[45] Woodhouse KD et al. A review of shared decision-making and patient decision aids in radiation oncology. J Cancer Educ 2017;32:238–45. https://doi.org/10.1007/s13187-017-1169-8.
[46] van Tol-Geerdink JJ et al. Do patients with localized prostate cancer treatment really want more aggressive treatment? J Clin Oncol 2006;24:4581–6. https://doi.org/10.1200/JCO.2006.05.9592.
[47] Serban M et al. Ring versus ovoids and intracavitary versus intracavitary-interstitial applicators in cervical cancer brachytherapy: Results from the EMBRACE I study. Int J Radiat Oncol Biol Phys 2020;106:1052–62. https://linkinghub.elsevier. com/retrieve/pii/S0360301619345328.
[48] Moore DH, Mendelsohn ML. Optimal treatment levels in cancer therapy. Cancer 1972;30:97–106. https://doi.org/10.1002/1097-0142(197207)30:1<97::AID-CNCR2820300116>3.0.CO;2-M.
[49] Robert Andrews, J. Benefit, risk, and optimization by roc analysis in cancer radiotherapy. *International Journal of Radiation Oncology*Biology*Physics* **11**, 1557–1562 (1985). URL https: //linkinghub.elsevier.com/retrieve/pii/0360301685903451.
[50] Hoffmann AL, Huizenga H, Kaanders JH. Employing the therapeutic operating characteristic (TOC) graph for individualised dose prescription. Radiat Oncol 2013;8:55. https://doi.org/10.1186/1748-717X-8-55.
[51] Moore KL. Automated radiotherapy treatment planning. Semin Radiat Oncol 2019;29:209–18. https://linkinghub.elsevier.com/retrieve/pii/S1053429619300128.