Received: 27 October 2021 | Revised: 1 April 2022 | Accepted: 3 April 2022

DOI: 10.1002/mp.15683

RESEARCH ARTICLE

MEDICAL PHYSICS

Development of a Monte Carlo based robustness calculation and evaluation tool

Hannes A. Loebner[1] | **Werner Volken**[1] | **Silvan Mueller**[1] | **Jenny Bertholet**[1] | **Paul-Henry Mackeprang**[1] | **Gian Guyer**[1] | **Daniel M. Aebersold**[1] | **Marco F. M. Stampanoni**[2] | **Peter Manser**[1] | **Michael K. Fix**[1]

[1]Division of Medical Radiation Physics and Department of Radiation Oncology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[2]Institute for Biomedical Engineering, ETH Zürich and PSI, Villigen, Switzerland

Correspondence
Hannes Anton Loebner, Abteilung fuer medizinische Strahlenphysik, Friedbuehlschulhaus, Universitaetsspital Bern, Freiburgstrasse 18, 3010 Bern, Switzerland.
Email: Hannes.Loebner@insel.ch

Funding information
Varian Medical Systems

Abstract

Background: Evaluating plan robustness is a key step in radiotherapy.

Purpose: To develop a flexible Monte Carlo (MC)-based robustness calculation and evaluation tool to assess and quantify dosimetric robustness of intensity-modulated radiotherapy (IMRT) treatment plans by exploring the impact of systematic and random uncertainties resulting from patient setup, patient anatomy changes, and mechanical limitations of machine components.

Methods: The robustness tool consists of two parts: the first part includes automated MC dose calculation of multiple user-defined uncertainty scenarios to populate a robustness space. An uncertainty scenario is defined by a certain combination of uncertainties in patient setup, rigid intrafraction motion and in mechanical steering of the following machine components: angles of gantry, collimator, table-yaw, table-pitch, table-roll, translational positions of jaws, multileaf-collimator (MLC) banks, and single MLC leaves. The Swiss Monte Carlo Plan (SMCP) is integrated in this tool to serve as the backbone for the MC dose calculations incorporating the uncertainties. The calculated dose distributions serve as input for the second part of the tool, handling the quantitative evaluation of the dosimetric impact of the uncertainties. A graphical user interface (GUI) is developed to simultaneously evaluate the uncertainty scenarios according to user-specified conditions based on dose-volume histogram (DVH) parameters, fast and exact gamma analysis, and dose differences. Additionally, a robustness index (RI) is introduced with the aim to simultaneously evaluate and condense dosimetric robustness against multiple uncertainties into one number. The RI is defined as the ratio of scenarios passing the conditions on the dose distributions. Weighting of the scenarios in the robustness space is possible to consider their likelihood of occurrence. The robustness tool is applied on IMRT, a volumetric modulated arc therapy (VMAT), a dynamic trajectory radiotherapy (DTRT), and a dynamic mixed beam radiotherapy (DYMBER) plan for a brain case to evaluate the robustness to uncertainties of gantry-, table-, collimator angle, MLC, and intrafraction motion. Additionally, the robustness of the IMRT, VMAT, and DTRT plan against patient setup uncertainties are compared. The robustness tool is validated by Delta4 measurements for scenarios including all uncertainty types available.

Results: The robustness tool performs simultaneous calculation of uncertainty scenarios, and the GUI enables their fast evaluation. For all evaluated plans and uncertainties, the planning target volume (PTV) margin prevented major clinical target volume (CTV) coverage deterioration (maximum observed standard deviation of $D98\%_{CTV}$ was 1.3 Gy). OARs close to the PTV experienced larger dosimetric deviations (maximum observed standard deviation of $D2\%_{chiasma}$ was 14.5 Gy). Robustness comparison by RI evaluation against patient setup uncertainties revealed better dosimetric robustness of the VMAT and DTRT plans as compared to the IMRT plan. Delta4 validation measurements agreed with calculations by >96% gamma-passing rate (3% global/2 mm).
Conclusions: The robustness tool was successfully implemented. Calculation and evaluation of uncertainty scenarios with the robustness tool were demonstrated on a brain case. Effects of patient and machine-specific uncertainties and the combination thereof on the dose distribution are evaluated in a user-friendly GUI to quantitatively assess and compare treatment plans and their robustness.

KEYWORDS
Monte Carlo, plan evaluation, robustness (to patient and machine-related uncertainties)

1 | INTRODUCTION

A key step in the radiotherapy treatment workflow is the treatment plan evaluation both in terms of dose parameters and in terms of robustness.[1] In recent years, evaluating robustness to uncertainties on the patient side (e.g., setup uncertainty[2]) or in the mechanical accuracy of the treatment machine (e.g., multileaf-collimator (MLC) leaf positioning accuracy[3]) has become an essential part of plan quality assessment.[4]

To quantify the impact of an uncertainty on the plan robustness, its influence on the dosimetric quality of the plan has to be determined. Due to interpatient anatomical variations and the large variety of treatment techniques available, it is difficult to make a general statement relating a specific type of uncertainty to a dosimetric consequence.

Treatment plan robustness depends on the applied technique. Particularly, the complexity of the techniques impedes the straightforward understanding of the dosimetric impact of a type of uncertainty. Developments in intensity-modulated treatment techniques for C-arm treatment units aim to improve plan quality by using increased degrees of freedom (DoF) and consequently increasing the plan complexity for the novel techniques. Volumetric modulated arc therapy (VMAT) increases the DoF by including dynamic gantry rotation as compared to intensity modulated radiotherapy (IMRT), and has become standard of care of radiotherapy.[5–7] Dynamic trajectory radiotherapy (DTRT), which extends VMAT by dynamic table and collimator rotations during delivery, adds two additional DoF, while ensuring acceptable delivery times.[8,9] An additional DoF, and additional complexity, is introduced in mixed beam radiotherapy (MBRT[10]) and dynamic MBRT (DYMBER[11]) by combining photon and electron beams to use the radiation type specific advantages of sharp beam penumbra (photons) and distal dose fall-off (electrons). Treatment plan robustness therefore depends on the treatment technique. For instance, uncertainties in gantry position have different effects on the dose distribution for the dynamic gantry rotation in VMAT as compared to the static gantry in IMRT, and patient setup uncertainties have different effects on photon treatment compared to electron treatments.

Usually, a differentiation between systematic and random uncertainties is made.[12] Uncertainties in patient setup have both systematic and random components. On the machine side, systematic uncertainties such as calibration errors, and intrafraction uncertainties of each machine component, influence the delivery accuracy and thus the delivered dose distribution. In recent years, the study of machine logfiles has been a focus topic in research. Logfiles contain time-resolved information of the machine status during delivery and are used to calculate the fraction-specific uncertainty of the machine. In routine clinical practice, logfiles are used to monitor the performance of treatment machines, improve efficiency in quality assurance (QA) workflows,[13–17] and examine the robustness of treatment plans.[18,19]

With this increasing variety and complexity in treatment techniques, it is not sufficient anymore to assess, for example, MLC positioning uncertainty[20,21] or uncertainties in patient setup[2] alone. These investigated uncertainty scenarios have to be extended to account for the added DoF of new techniques and the uncertainties must be assessed in combination to achieve a more comprehensive robustness assessment of a treatment plan. There is a need for simple and comprehensive calculation and evaluation tools of the dosimetric impact of

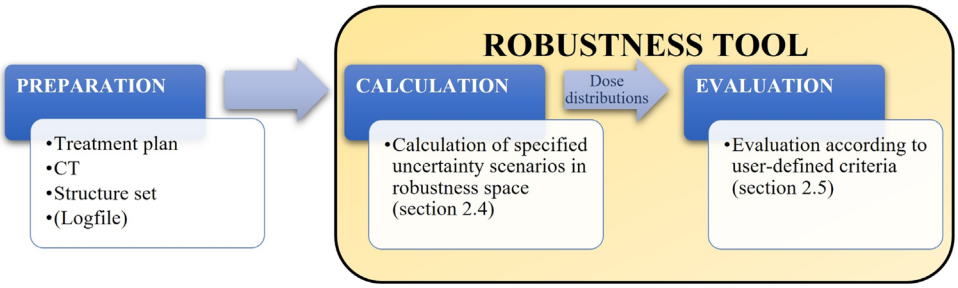

FIGURE 1 Workflow for the robustness tool. *Preparation*: The treatment plan is created in a treatment planning system (TPS) based on the CT and structure set. To include logfile and time-resolved information, the created treatment plan must be delivered first and the logfile recorded. *Part 1, Calculation*: user defines the robustness space in terms of uncertainty scenarios. Subsequently, the respective dose distributions incorporating the desired uncertainties are calculated (see Section 2.4). *Part 2, Evaluation*: evaluation of the dose distributions of the robustness space (see Section 2.5)

all patient and machine-related uncertainties, including their interplay.

Therefore, this work aims to develop a robustness tool to evaluate the dosimetric robustness of treatment plans. The robustness tool will be applicable to a wide range of treatment techniques and used to investigate the impact of patient and machine component-related uncertainties on the dose distribution, individually and in combination. To ensure the accurate calculation of the dose distributions, including the aforementioned uncertainties, Monte Carlo (MC) dose calculation[22] using the Swiss Monte Carlo Plan (SMCP)[23,24] will serve as the backbone of this tool.

Finally, following the idea of a plan quality index,[25] we introduce a robustness index (RI), which condenses dosimetric robustness of multiple uncertainties into one number. This streamlines the robustness evaluation and robustness comparison of different treatment plans, for the considered uncertainties.

2 | MATERIALS AND METHODS

The robustness tool consists of two components: the calculation and the evaluation part. In the calculation part, the dose distribution including the desired uncertainties is calculated. In the evaluation part, the dosimetric impact of the uncertainties on the dose distribution is assessed. Dose calculation uses SMCP on a high-performance computing cluster for efficient calculation.[24] To evaluate the dosimetric impact of the uncertainties, a GUI is developed. The workflow is illustrated in Figure 1.

2.1 | Terminology

To assess robustness of a treatment plan, a so-called *robustness space* is evaluated. The robustness space has the dimensionality N, related to the N uncertainty types considered for the evaluation. The robustness space is spanned up by N *uncertainty axes* $a_i, i \in \{1 \dots N\}$, with z_i *uncertainty scenarios* at $\{n^1_{a_i}, n^2_{a_i}, \dots, n^{z_i}_{a_i}\}$ along the axis a_i. The location of a single scenario in the robustness space is given by $n^{x\in\{1\dots z_1\},\dots,x\in\{1\dots z_N\}}_{a_1,\dots,a_N}$. Thus, each scenario represents a combination of uncertainties. An uncertainty axis is characterized by the origin of the considered uncertainty: it can be patient (e.g., setup, motion) or machine (e.g., gantry uncertainty) related. The uncertainty axes and scenarios for calculation and evaluation are specified by the user of the robustness tool. A scenario is associated with the dose distribution including the corresponding uncertainties. In the center of the robustness space, the *reference scenario* represents the nominal plan with the planned dose distribution involving no uncertainty, the reference CT, and the reference structure set. For example, the robustness space considering patient setup uncertainties with respective scenarios in longitudinal, lateral, and vertical patient setup direction and their combination, has scenarios in three dimensions.

The following uncertainty axes are implemented in the tool: patient setup uncertainties (longitudinal, lateral, vertical, table rotation angle, table pitch angle, table roll angle) are modeled by table translations and rotations. Further uncertainty axes consider the mechanical accuracy of machine components (gantry angle, collimator angle, MLC leaf positions, translational jaws position), monitor units, and rigid three DoF intrafraction patient motions in longitudinal, lateral, and vertical direction. Additionally, the tool supports the re-calculation of the plan on different CTs to account for anatomical changes. If requested, each scenario has its own CT and structure set. Further not yet considered uncertainty axes are easily integrated in the tool.

2.2 | Data structure

Scenarios in the robustness space are structured according to two options: in the *gridded data* (GD) option, where the scenarios are ordered in a regular grid. Thus,

the dose distributions of all possible combinations of the considered uncertainty axes are calculated: Together, the N uncertainty axes define a set of $z_1 \times z_2 \times ... \times z_N$ scenarios in the robustness space. The *spotted data* (SD) option requires the specification of each individual uncertainty scenario at $n_{a_1,...,a_N}^{x\in\{1...z_1\},...,x\in\{1...z_N\}}$, for calculation of the respective dose distribution. The SD option potentially limits the number of calculations and subsequent evaluations, which becomes particularly useful to probe the robustness space, for example, for the most sensitive uncertainty axes.

2.3 | Uncertainty type

The robustness tool enables to calculate and evaluate patient and machine-related uncertainties. For the purpose of patient-related setup uncertainties, a Gaussian distribution is used.[12] Thereby the mean and the standard deviation (sigma) represent the systematic and the random component of the uncertainty. During the Monte Carlo dose calculation, a shift is sampled from the Gaussian (mean, sigma) and applied to the particle exiting the linac head. Thus, for each particle exiting the linac head, such a shift is applied following the Gaussian distribution. As this distribution is applied on a particle-by-particle basis, there is no need to simulate several fractions for a specific setup uncertainty. The obtained dose distribution directly accounts for the setup uncertainty distribution. This approach saves substantial computational resources but assumes a sufficient number of fractions (>10), as shown by van Herk et al.[26] In the context of robust optimization, the effect of number of fractions on the simulation of random setup uncertainties and robustness has been investigated previously by Fredriksson.[27]

For machine-related uncertainties, the robustness tool models global systematic uncertainties by constant offsets in the respective machine components. To simulate treatment plan and machine-specific uncertainty combinations, information of the machine logfile is used. To this end, the plan has to be delivered at least once prior to application to record the logfile, which, for example, could be done during pretreatment quality assurance. In the logfile of a TrueBeam (Varian Medical Systems, Palo Alto, CA, USA), the status of all machine components is logged as *expected* and *actual* value during delivery at a rate of 50 Hz with absolute time stamp. The difference between expected and actual for each control point can be assigned as control point-specific local systematic uncertainty. The robustness tool enables to consider and to scale these uncertainties for the subsequent respective dose calculation.

The time-resolved machine logfile information can be further used to synchronize a time series of intrafraction patient-related motion (e.g., breathing) with machine parameters.

2.4 | Calculation

The reference treatment plan is generated with an Eclipse research version (Varian Medical Systems, Palo Alto, CA, USA). An input file, defining the treatment machine axis at all control points of the treatment for the nominal scenario, is created. The user specifies the desired robustness space, and the input file is automatically adapted accordingly using a python framework to serve as an input for the subsequent calculation of the dose distributions of the scenarios in the robustness space, using SMCP.[23,24,28] Calculation voxel size is $0.25 \times 0.25 \times 0.25\,\mathrm{cm}^3$. The statistical uncertainty of the Monte Carlo calculated dose distributions presented in this work is <1.2% (one standard deviation). The number of simulated primary particles per calculation is in the order of 10^8.

2.5 | Evaluation

Evaluation of the uncertainty scenarios dose distributions is performed according to *DVHs*, *gamma passing rate* and *dose differences*. To provide a flexible and user-friendly robustness evaluation of treatment plans, a GUI is developed. The read-in of the input of the GUI, loading of the robustness space data (dose distributions, CTs, structures), and the calculation of DVHs and gamma passing rate is parallelized on the number of available CPU cores. The computer memory required for the evaluation is roughly proportional to the size of the dose distribution times the number of scenarios. The GUI is implemented in C++ utilizing the Qt toolkit[29] and has a dynamic structure.

The user of the robustness tool interactively defines the evaluation conditions, including thresholds for acceptance to evaluate the robustness of a scenario in the robustness space. Possible conditions are based on $Dx\%$, $Vx\%$, D_{mean} or gamma passing rate, evaluated for selected structures. To assess the "distance" of the individual quantity to the specified acceptable threshold, a *conditions meter* c is introduced as follows:

$$c = \frac{Q_{\mathrm{current}} - Q_{\mathrm{ref}}}{Q_{\mathrm{acceptable}} - Q_{\mathrm{ref}}} \tag{1}$$

where c is the ordinate for a given quantity Q (e.g., D_{mean}). Q_{current} and Q_{ref} are the parameter values for the current and reference scenario, respectively. $Q_{\mathrm{acceptable}}$ is the threshold for the parameter value used to determine if Q is robust. For $c > 1$, the

scenario is not considered robust and for $c < 1$ it is robust. For $c = 0$, the scenario fulfils the condition equally well as the reference scenario and for $c < 0$ the scenario achieves the condition better than the reference scenario.

The calculation of the gamma passing rate is based on the algorithm of Ju et al.,[30] which calculates the gamma without using a linear interpolator and discretization of the search space (Supporting Material S1). This has the advantage that the time needed for calculating the gamma passing rate is independent of the distance- and dose difference criteria and an exact gamma value is calculated: the common limitation of lacking an infinite resolution grid is overcome by calculating the gamma passing rate values of each reference dose point as the closest geometric distance between this point to the hypersurface defined by the evaluation dose distribution. Additionally, a maximal gamma threshold value, representing the maximal difference in gamma value to the nominal scenario is set, after which the calculation is automatically terminated to save calculation time. The robustness tool employs a global gamma analysis, with the reference scenario set as the reference dose in each evaluation. The dose criterion is given relative to a user-defined value. The default is the prescribed dose. Additionally, a low-dose threshold relative to the beforementioned value is used.

Visual evaluation of the difference between dose distributions of any scenarios in the robustness space is conducted in three dimensions (lateral, sagittal, coronal), with an adjustable dose difference threshold.

2.6 | Interpolation between scenarios

Evaluation of the treatment plan robustness is not limited to the initially defined uncertainty scenarios. The scenarios populating the robustness space serve as sampling points, between which interpolation is applied. To this end, a metric defining the distance d between two scenarios at b and c in the robustness space is specified as

$$d^2 = \sum_{l=1}^{N} \left(\frac{b_l - c_l}{m_l} \right)^2 \tag{2}$$

with the uncertainty normalization $M = (m_1, m_2, \ldots, m_N)$. The uncertainty normalization is defined by default as

$$m_l = \frac{n_{a_l}^{z_l} - n_{a_l}^{1}}{z_l - 1} \tag{3}$$

but is adjustable. The resulting distance is unitless. The evaluation quantity Q can then be interpolated based on the nearest neighbors.

2.7 | Robustness index

To summarize the information of treatment plan robustness considering multiple uncertainties, the robustness index (RI) is introduced. When all scenarios in the robustness space are considered equally important, the RI describes the fraction of scenarios passing the conditions: $\mathrm{RI} = \frac{N_\mathrm{p}}{N_\mathrm{all}}$, where N_p is the number of scenarios passing all robustness conditions and N_all is the number of all scenarios in the present robustness space.

However, a greater distance to the location of the reference scenario usually represents a decrease in the likelihood of occurrence of such a scenario. Therefore, a second option to calculate RI is introduced, given by $\mathrm{RI} = \frac{\sum_{t=1}^{N_\mathrm{p}} w_t}{\sum_{t=1}^{N_\mathrm{all}} w_t}$ with $w_t = e^{\frac{-d_t^2}{2R_\mathrm{E}^2}}$. d_t is the distance between the location of scenario t and the reference scenario in the robustness space and R_E is the *evaluation range*, the standard deviation of this scenario distribution. The maximal value of the RI is 100% (in the weighted and unweighted case) and states that all scenarios pass the conditions. In the unweighted case, an RI of 10% means that 10% of the scenarios in the robustness space pass the conditions. In the weighted case, an RI of 10% describes that the fraction of the summed weights of the passing scenarios is 10%. Thereby the scenarios are weighted according to a Gaussian distribution $G(\mu = 0, \sigma = R_\mathrm{E})$ centered around the reference scenario (Supporting Material S2). The mean and standard deviation of a quantity Q evaluated on the whole robustness space change to a weighted *mean*, $\mu_Q = \frac{\sum_t w_t Q_t}{\sum_t w_t}$ and weighted *standard deviation*, Σ_Q with $\Sigma_Q^2 = \frac{\sum_t w_t Q_t^2 \sum_t w_t - (\sum_t w_t Q_t)^2}{(\sum_t w_t)^2 - \sum_t w_t^2}$, where w_t is the Gaussian weighted distance factor from the reference as defined above and Q_t is the value of the quantity Q at scenario t.

2.8 | Application and demonstration of the robustness tool

Application of the robustness tool and the GUI is demonstrated on different treatment plans for a right-sided brain case (Figure 2), namely, an IMRT, a VMAT, a DTRT, and a DYMBER plan (Table 1). Prescribed dose is 60 Gy to 50% of the planning target volume (PTV). The clinical target volume CTV–PTV margin is 0.5 cm. The respective plans are optimized according to clinical goals applied in our institute. Application of the main functionalities of the GUI is demonstrated on the DTRT plan. To conduct the robustness comparison against uncertainties in patient setup of different treatment techniques, applications 1–6 (Table 1) are assessed. To demonstrate the flexibility of the robustness tool in terms of different

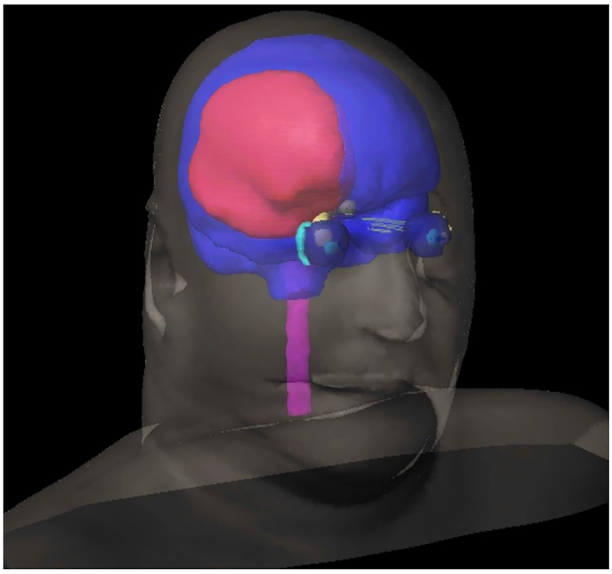

FIGURE 2 Right-sided glioblastoma brain case. The PTV is shown in red, the OARs in color

treatment techniques and uncertainty scenarios, applications 7–12 are investigated.

The robustness conditions are set to: $D98\%$ and $D2\%$ of the CTV are not allowed to be reduced or increased by more than 1 Gy, respectively. Additionally, $D2\%$ and mean dose are not allowed to be increased by more than 1 Gy for serial and parallel OARs, respectively. The plan robustness is then quantified by the RI: once for all scenarios weighted equally and once with $R_E = 1.5$ to increase the weight of the more likely scenarios.

2.9 | Validation of the robustness tool

Extensive validation of the SMCP dose calculation framework has been conducted in the past.[9–11,24,28,31] This included validation of photon-based treatment techniques such as IMRT, VMAT, DTRT, as well as mixed beam treatment techniques, such as DYMBER. Measurements are conducted with the Delta4 device (ScandiDos, Uppsala, Sweden). First, 10×10 cm^2 fields and 5×5 cm^2 fields at different gantry angles ($0°, 90°, 180°$) are delivered and the measured dose is compared to the calculation. Second, to validate the accuracy of the dose calculation of selected scenarios (Table 2) in the robustness space of the VMAT plan for the brain case (Figure 2), the scenarios are calculated with the robustness tool and the corresponding XML files needed for delivery on a TrueBeam (Varian Medical Systems, Palo Alto, CA, USA) in developer mode are created, the plans are delivered, and the delivered dose is measured and compared against the calculation.

To validate the calculation of quantities Q, for example, DVH parameters such as volume, $Dx\%$ and D_{mean}, the robustness tool calculations are compared to the results of the corresponding Eclipse implementation. The dose distribution of the reference VMAT plan is loaded into Eclipse, and the DVH parameters for several organs are compared to the evaluation returned by the robustness tool. The gamma passing rate calculations of the robustness tool are validated against the open-source gamma passing rate calculation of pymedphys 0.37.1, an implementation based on the work of Wendling et al.[32]: gamma passing rates of the reference scenario of the VMAT plan and the scenarios in the robustness space incorporating systematic setup uncertainties in lateral, longitudinal, and vertical directions between -0.5 and 0.5 cm are calculated and compared.

3 | RESULTS

3.1 | GUI for evaluation

Robustness evaluation with the GUI is shown in Figure 3 for the DTRT plan Table 1, application 8. The DVH viewer (Figure 3, number 1) represents the DVH distributions of the selected structures over all scenarios in the robustness space as DVH bands. The boundaries of the DVH bands correspond to the scenarios, which induce the maximal deviations in terms of dose distribution for this structure. The reference DVH is indicated by a solid line and the DVH of the current scenario is shown with a dashed line. In the axes selection window (Figure 3, number 3), the user switches the uncertainty axes of the robustness space to select a two-dimensional plane in the robustness space. The selected plane is displayed in the robustness map (Figure 3, number 2) in real time. Additionally, the evaluation range is superimposed on the robustness map: in Figure 3, an R_E of 1.5 and standard metric is chosen to emphasize the smaller uncertainties. The evaluation range is visualized by the shaded area. The conditions meter for the selected conditions set in the conditions list is shown on the bottom right of Figure 3.

3.1.1 | Conditions list

The conditions from Section 2.8 are specified in the conditions list (Figure 3, number 5 opens the conditions list seen in Figure 4) and are adjustable on the fly according to the interest of the user. The following structures are part of the evaluation: brain, brainstem, left and right eye (Eye_l./r.), chiasma, left and right lens (Lense_l./r.), left and right optic nerve (N.opticus_l./r.), left and right lacrimal glands (Lacrimal_l./r.), CTV, and PTV.

TABLE 1 Application of the robustness tool, demonstrated for different treatment techniques and uncertainties

#Application	Treatment technique	Description of uncertainty axes and scenarios	#Scenarios
1	IMRT (3 fields)	Systematic patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(\mu, 0)$, $\{\mu = -0.5, -0.2, 0.0, 0.2, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
2	IMRT (3 fields)	Random patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(0, \sigma)$, $\{\sigma = 0.0, 0.1, 0.2, 0.3, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
3	VMAT (2 arcs)	Systematic patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(\mu, 0)$, $\{\mu = -0.5, -0.2, 0.0, 0.2, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
4	VMAT (2 arcs)	Random patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(0, \sigma)$, $\{\sigma = 0.0, 0.1, 0.2, 0.3, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
5	DTRT (3 trajectories)	Systematic patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(\mu, 0)$, $\{\mu = -0.5, -0.2, 0.0, 0.2, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
6	DTRT (3 trajectories)	Random patient setup uncertainty: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(0, \sigma)$, $\{\sigma = 0.0, 0.1, 0.2, 0.3, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	125
7	IMRT (3 fields)	Global systematic uncertainty counter-moving jaws X1 and X2 and Y1 and Y2, {−1, −0.5, 0.0, 0.5, 1.0 cm}; Global systematic uncertainty leaf bank A and B, {−0.1, −0.05, 0.0, 0.05, 0.1 cm}; and combinations of them	625
8	DTRT (3 trajectories)	Global systematic uncertainty in gantry-, table-, collimator angle {−4.0°, −2.0°, −1.0°, 0.0°, 1.0°, 2.0°, 4.0°}, and combinations of them	343
9	DYMBER (3 photon trajectories and 1 electron field [6, 9, 12, 15, 18, 22 MeV])	Systematic patient setup uncertainty for electron fields: longitudinal, lateral, and vertical directions modeled with Gaussian distribution $G(\mu, 0)$, $\{\mu = -0.5, 0.0, 0.5\,\text{cm}\}$ ∀ axes, and combinations of them	27
10	VMAT (2 arcs)	Rigid intrafraction motion in transversal and sagittal planes, Rotation amplitude of isocenter around dens axis in transversal and sagittal planes, {0.0°, 0.5°, 1.0°, 2.0°, 3.0°}, and combinations of them	25
11	VMAT (2 arcs)	Global systematic collimator and table angle uncertainty, {−4.0°, −2.0°, −1.0°, 0.0°, 1.0°, 2.0°, 4.0°}; Local systematic gantry uncertainty: mean absolute gantry logfile difference between expected and actual scaled to {0.0°, 1.0°, 2.0°, 3.0°, 4.0°}, and combinations of them	245
12	VMAT (2 arcs)	Global systematic uncertainty in table angle and gantry angle, {−1.0°, 0.0°, 1.0°}; Local systematic gantry uncertainty: mean absolute gantry logfile difference between expected and actual scaled to {0.0°, 1.0°}; Systematic and random lateral patient setup uncertainty: modeled with Gaussian distribution $G(\mu, \sigma)$, $\{\mu = -1.0, 0.0, 1.0\,\text{cm}; \sigma = 0.0, 0.5\,\text{cm}\}$ and combinations of them	108

Note: Applications 1–6 evaluate systematic and random patient setup uncertainty, and applications 7–12 evaluate different robustness spaces for different treatment techniques.

TABLE 2 Validation scenarios

Validation scenario	Description
1	10×10 cm^2 field at gantry angles 0°, 90°, 180°
2	5×5 cm^2 field at gantry angles 0°, 90°, 180°
3	Reference scenario (no uncertainty)
4	Systematic patient setup uncertainty of +1 cm in lateral, longitudinal, and vertical directions
5	Systematic uncertainty of 1° in gantry
6	Systematic uncertainty of 1° in gantry and collimator
7	Systematic uncertainty of 1° in gantry, collimator, and table rotation
8	Systematic uncertainty of 1 mm for MLC leaf 40 in leafbank B
9	Systematic uncertainty of 3 cm in Jaw X1
10	Intrafraction motion of 1 cm in longitudinal direction during treatment
11	Systematic 10% more MU
12	Gantry logfile uncertainty scaled up by factor 10
13	Systematic gantry uncertainty of 1° combined with systematic patient setup uncertainty of +1 cm in lateral, longitudinal, and vertical directions

Note: Validation scenarios 1–2: validation of 10×10 and 5×5 cm^2 open fields at different gantry angles. Validation scenarios 3–13: specific uncertainty scenarios for validation of the robustness space of the VMAT plan.

3.1.2 | Conditions meter

The conditions meter in Figure 3, number 5, displays how well the current scenario is fulfilling the user-defined conditions. If all conditions are passed, the respective rectangle in the robustness map (Figure 3, number 2) is marked green, and red, if one condition is not fulfilled. To evaluate changes in the distance or dose difference criteria in the calculation of gamma passing rate, the gamma passing rate calculation needs to be restarted by ticking the red circle in the main GUI window (Figure 3, number 4).

3.1.3 | Dose window

The dose window (Figure 5) displays the dose distributions of the selected scenario (systematic uncertainty of 1° in table and gantry angle: Figure 5, number 1), the reference scenario (Figure 5, number 2), and the dose difference (Figure 5, number 3) superimposed on the respective CT. The dose difference is always superimposed on the CT of the reference scenario and the structures of the underlying anatomy are visualized by their contours. The user explores the dose distributions in the transversal, coronal, and sagittal planes by scrolling through them, and switches from one plane to the other with the help of the crosshairs. Additionally, an adjustable dose threshold is implemented to visualize dose differences in a desired range.

3.1.4 | Statistics window

The statistics window, displayed in Figure 6, summarizes dosimetric key quantities for all OARs. Additionally, the RI for the selected evaluation range is displayed at the bottom.

3.1.5 | Interpolation

A robustness map of SD and corresponding GD is shown in Figure 7. Here only some scenarios are available (shown as white dots). Green and red areas indicate whether a scenario is considered robust and acceptable or not, respectively. Between the available scenarios, the robustness map displays gray rectangles (Figure 7, number 1). So far, no information on the robustness of these areas in the robustness space is available. These empty spaces are filled by interpolation. Besides interpolation, the *zoom* functionality is implemented in the GUI. It inspects a single two-dimensional slice through the robustness space in greater detail. The zoom function increases the number of scenarios along the uncertainty axes of this slice by a factor of two, thus leading to additional combinations of these uncertainty axes in the selected plane and performs an interpolation on all inserted scenarios. Interpolation and zooming are implemented on a multithreaded basis.

3.2 | Robustness evaluation of different treatment plans

The results of the applications 1–12 from Table 1 are shown in Table 3. The default uncertainty normalization is chosen. Furthermore, the standard deviation Σ of the $D98\%_{CTV}$ and $D2\%_{CTV}$ as well as the structure with the greatest standard deviation in mean dose and in $D2\%$ are displayed. RI is higher for a weighted analysis ($R_E = 1.5$), compared to equal weighting of the scenarios. Due to the Gaussian weighted evaluation method, RI increases, as scenarios located closer to the reference have more weight compared to scenarios located further away from the reference scenario and they usually pass the conditions better.

Evaluating applications 1–6 compares the robustness of different treatment techniques to patient setup uncertainties. The RI indicates better dosimeteric robustness to setup uncertainties for the VMAT and the DTRT plans as compared to the IMRT plan. Assessing RI and the standard deviation of $D98\%_{CTV}$ and $D2\%_{CTV}$ show that the CTV–PTV margin compensates for the uncertainties to maintain CTV coverage, but some of the OARs

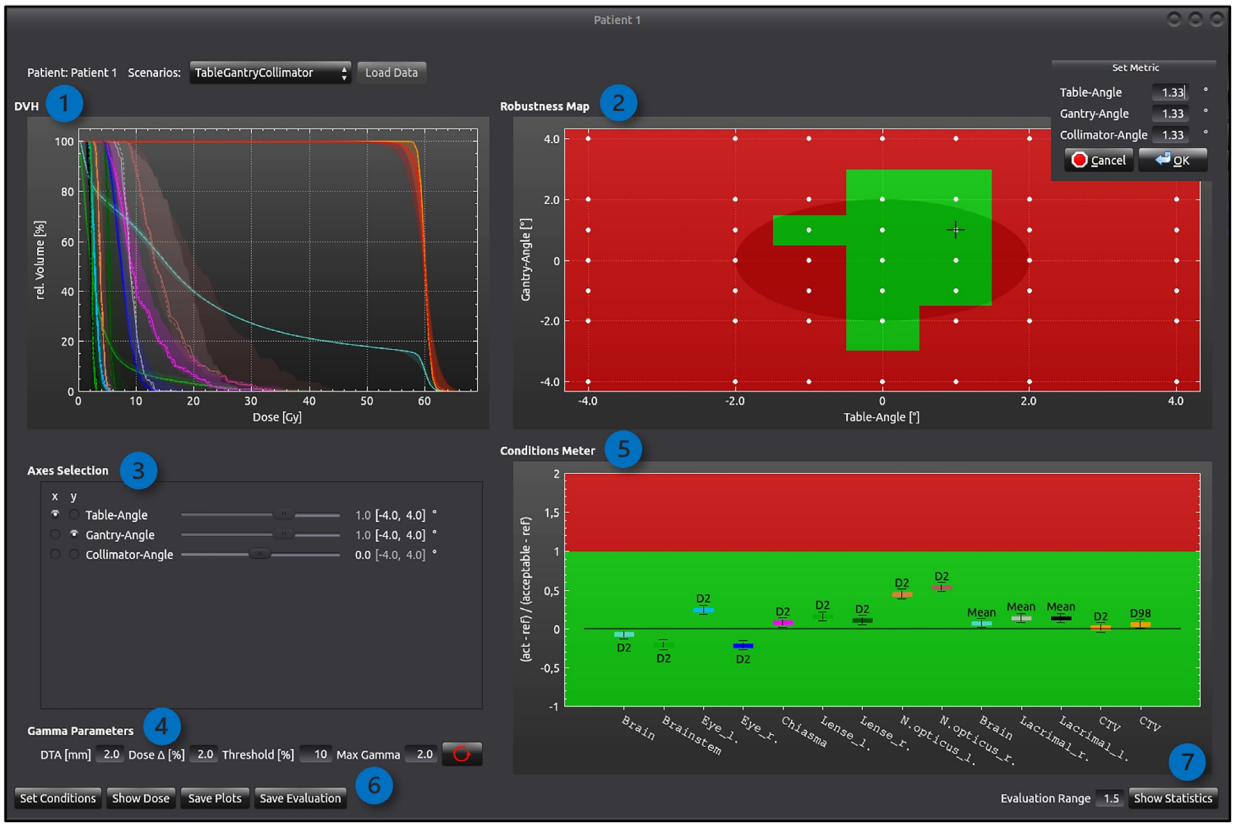

FIGURE 3 Main GUI of the robustness evaluation tool. Top left, 1: *DVH viewer* with DVH bands of all scenarios of the robustness space of the DTRT plan (Table 2, application 8). Top right, 2: *Robustness map* displaying 2D plane of the robustness space and pop-up metric window to change the metric if needed, Bottom left, 3: *Axes selection* window to select a plane in the robustness space for closer inspection. Bottom left, 4: *Specification* of parameters *gamma passing rate* calculation. Bottom right, 5: *Conditions meter* for evaluating structure-specific conditions; 6 and 7: Open pop-up windows: *conditions list* (definition of robustness conditions), *dose window* (shows dose distribution and dose difference), and *statistics window* (summarizes key quantities of the robustness space)

are strongly affected by setup uncertainties, potentially leading to reject the plan completely when clinical constraints are no longer met.

In Figure 8, the DVH bands of the robustness space including random setup uncertainties of the IMRT, VMAT, and DTRT plans are shown. Especially the OARs near the PTV experience a great variation in their dose, as seen by the width of the DVH bands compared to the ones of PTV and CTV.

3.3 | Validation

In Table 4, the dose measurements of the selected scenarios (Table 2) are compared against the calculation. For the gamma analysis, dose difference/distance criteria of 2% (global)/1 mm for validation scenarios 1–2 and 3% (global)/2 mm for validation scenarios 3–13, including a 20% low-dose threshold, are applied in the evaluation for all measurements. All measurements agree with the calculation of the robustness tool with a gamma passing rate of >96.4%.

In Table 5, the results of the Eclipse implementation to calculate volume, $Dx\%$ and D_{mean} agree with the calculation of the robustness tool (e.g., PTV volume differentiates by 0.2%). For small structures greater differences (e.g., chiasma volume) occur. To validate the accuracy of the gamma passing rate calculation, a total of 64 scenarios are evaluated (2% of prescribed dose/2 mm, 20% threshold). The results cover a range of 91%–100% passing rate. The robustness tool and pymedphys deviate by a maximum of 0.8%. However, pymedphys needs substantially longer calculation time for the same number of CPU cores (on average: 5 min as compared to 2 s for the robustness tool).

4 | DISCUSSION

In this work, a novel robustness tool is developed and implemented to evaluate the robustness of treatment plans of different techniques to patient and machine-related uncertainties. In contrast to usually applied robustness evaluation procedures (e.g., evaluation limited to consider only setup uncertainties[2]), this tool has a great flexibility in terms of accurately calculating and evaluating different uncertainty scenarios and types (individually and in combination), in terms of

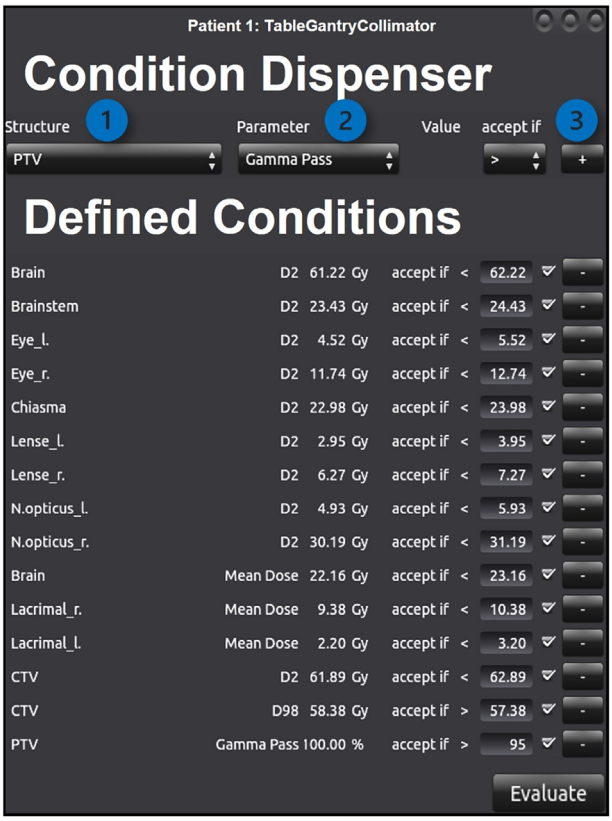

FIGURE 4 Conditions list opened by Figure 3, number 6. 1: Select structure, 2: Select parameter, 3: Add to evaluation

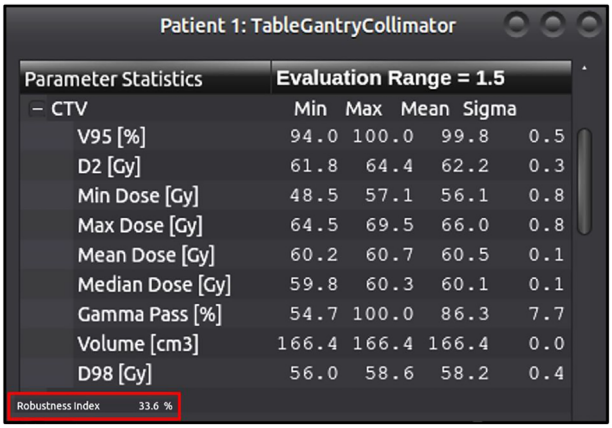

FIGURE 6 Statistics window (opened by Figure 3, number 7) displays key quantities and robustness index (RI) for all selected structures in the current evaluation range

simultaneously evaluating multiple uncertainty scenarios according to various criteria for the target and the OARs, and in terms of its applicability to different treatment techniques.

Accurate dose calculation of uncertainty scenarios is ensured by employing validated MC dose calculation with SMCP. It is necessary to simulate a sufficient number of primary particles for each MC dose calculation to achieve a reasonable statistical uncertainty[33] and hence reliable robustness results within the statistical uncertainty of the Monte Carlo calculated dose distributions. As MC dose calculations are usually computationally expensive, the tool is constructed in a modular way, and the dose calculation algorithm is interchangeable.

The robustness tool provides high flexibility to the user to enter specific scenarios for the robustness space. Owing to this flexibility in the calculation and evaluation of the robustness space, the developed robustness

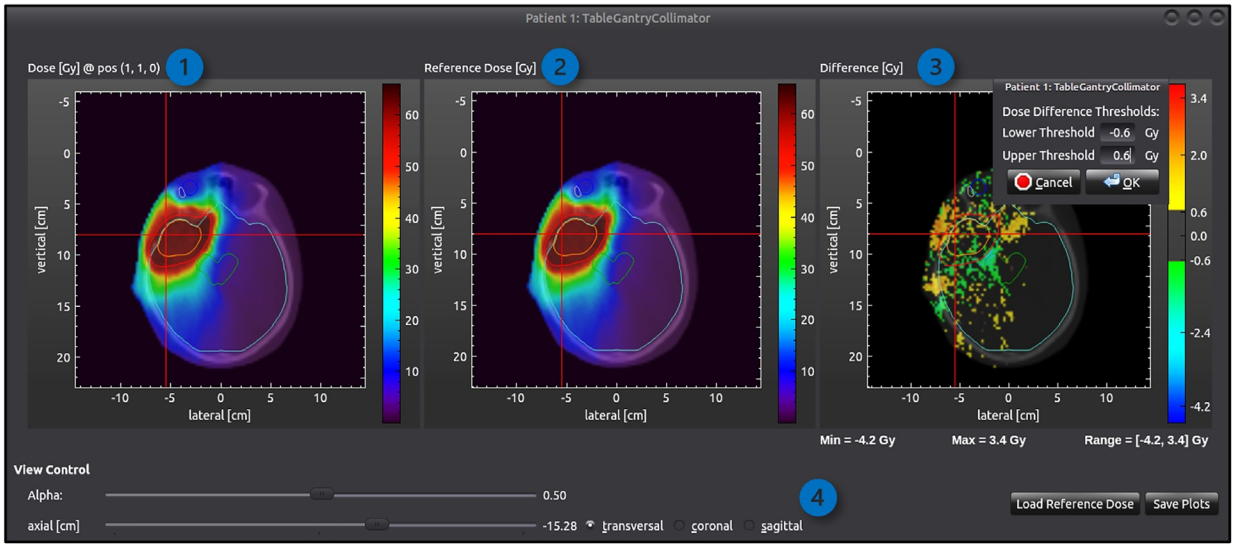

FIGURE 5 The dose window (opened by Figure 3, number 6) displays dose distributions superimposed on the CT. Structures are indicated by the fine lines (PTV and CTV in red and orange here). 1: Dose of current scenario. 2: Reference dose. 3: Dose difference between reference and current doses, includes a user-defined threshold to visualize relevant dose ranges. 4: View control to switch through transversal, coronal, and sagittal planes. The red cross serves as a guideline when switching planes from transversal (shown here) to coronal or sagittal

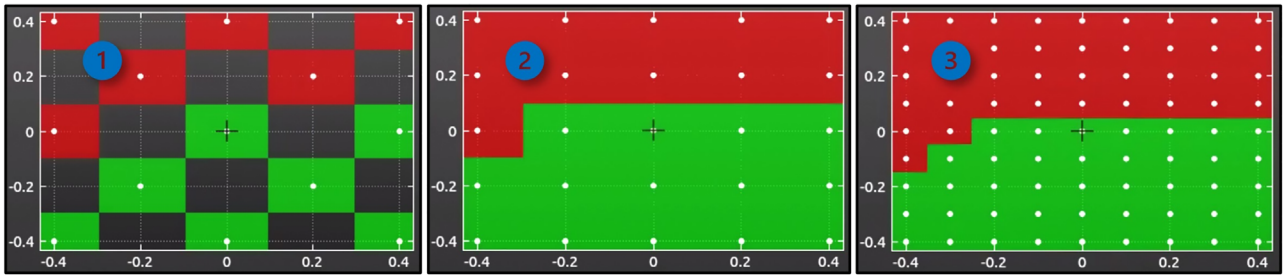

FIGURE 7 1: Robustness map of robustness space with SD, with scenarios at every other combination (white dots). 2: Robustness map of robustness space with GD. 3: Zoom functionality: doubles resolution of scenarios in this slice and interpolates on all inserted scenarios

TABLE 3 Evaluation of applications 1–12

#Application description	RI (all scenarios equally weighted)	RI (R_E1.5)	Σ of $D98\%_{CTV}$, Σ [Gy]	Σ of $D2\%_{CTV}$, Σ [Gy]	Structure with greatest Σ for mean dose, Σ [Gy]	Structure with greatest Σ for *D*2%, Σ Gy]
1. IMRT setup (systematic)	0.8%	2.2%	1.3	0.5	Optic nerve r.: 8.8	Optic nerve r.: 15.1
2. IMRT setup (random)	9.6%	31.3%	0.1	0.1	Chiasma: 1.8	Eye_r.: 2.4
3. VMAT setup (systematic)	8.0%	12.5%	0.5	0.3	Optic nerve r.: 7.3	Optic nerve r.: 14.4
4. VMAT setup (random)	14.4%	49.5%	0.1	0.2	Chiasma: 2.2	Chiasma: 1.8
5. DTRT setup (systematic)	16.8%	23.2%	0.5	0.4	Optic nerve r.: 6.2	Chiasma: 14.5
6. DTRT setup (random)	5.6%	23.4%	0.1	0.1	Optic nerve r.: 0.8	Chiasma: 1.5
7. IMRT MLC/Jaws	21.0%	29.9%	1.7	1.7	PTV: 1.9	Lacrimal gland r.: 1.9
8. DTRT gantry, table, collimator	12.0%	33.5%	0.4	0.3	Lacrimal gland r.: 0.9	Optic nerve r.: 1.5
9. DYMBER electron setup	25.9%	31.2%	0.9	0.7	Lacrimal gland r.: 0.4	PTV: 0.7
10. VMAT rigid intrafraction motion	16.0%	43.2%	0.2	0.8	Chiasma: 0.6	Optic nerve r.: 1.2
11. VMAT, logfile	5.1%	15.3%	0.2	0.2	Chiasma: 1.6	Lacrimal gland l.: 3.5
12. VMAT, systematic and random patient and machine uncertainties	13.9%	16.9	0.1	0.2	Optic nerve r.: 1.6	Chiasma: 3.6

Note: Robustness index (RI) for all scenarios weighted equally and R_E = 1.5. Standard deviations for $D98\%_{CTV}$ and $D2\%_{CTV}$ and structure with greatest standard deviation in mean dose and *D*2%. The calculation of standard deviations is based on the default metric and R_E = 1.5.

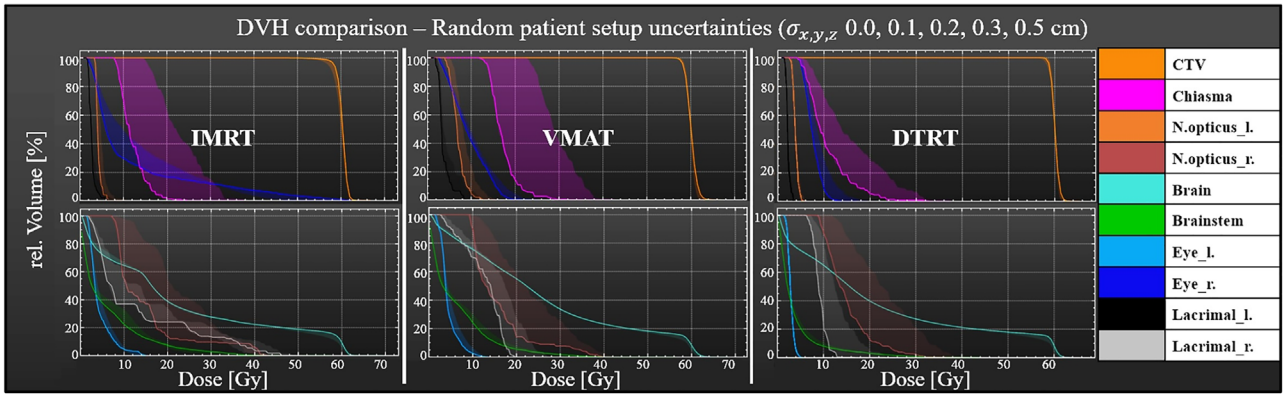

FIGURE 8 DVH bands for the IMRT, VMAT, and DTRT plans, including random uncertainties $G(0, \sigma_{x,y,z} = 0, 0.1, 0.2, 0.3, 0.5$ cm), in patient setup

TABLE 4 Gamma passing rate evaluation of scenarios specified in Table 2

Validation scenario	Gamma passing rate (2% global/1 mm)	
1	Average 98.5%	
2	Average 99.0%	
	Gamma passing rate (3% global/2 mm), Arc 1	Gamma passing rate (3% global/2 mm), Arc 2
3	97.8%	99.6%
4	98.7%	97.8%
5	97.0%	99.8%
6	96.4%	99.6%
7	97.0%	99.6%
8	97.2%	99.8%
9	98.7%	99.6%
10	99.8%	100%
11	97.4%	99.4%
12	99.0%	99.4%
13	99.4%	100%

Note: 2% (global)/1 mm dose difference and distance criteria, including a 20% dose threshold was applied to evaluate scenarios 1–2. The clinical dose difference and distance criteria, 3% (global)/2 mm, including a 20% dose threshold was applied for the evaluation of scenarios 3–13.

tool efficiently overcomes the restriction of standardized uncertainty scenarios: for example, limited robustness spaces, such as only considering MLC[34] or only patient setup[2,35] uncertainties. The user of the robustness tool defines the robustness space incorporating a variety of uncertainty types and their combinations. With the help of this tool, the dosimetric impact of systematic and random uncertainties in patient setup and uncertainties in the mechanical accuracy of machine components on the reference scenario are quantified. The robustness tool offers the possibility to simulate systematic and random setup uncertainties of Gaussian distributions. This reflects that, for example, the user can determine up to what level of uncertainty the given treatment plan is robust for the conditions considered, which then trigger, for example, an appropriate setup strategy. Additionally, it incorporates machine logfile information for realistic fraction-specific modeling of the uncertainties in each machine component, which has been determined to be a useful tool to assess machine-specific delivery uncertainties.[19,36–38] The robustness tool therefore permits comprehensive robustness evaluation adaptable to different use cases.

The flexibility to evaluate different treatment techniques, as for example described by Quian et al.,[36] is fundamental to the design of the robustness tool. Owing to the availability of validated dose calculation for various treatment techniques, robustness comparison of different treatment techniques, including different radiation sources, is possible with the robustness tool. This has the potential to redefine and restructure the treatment planning process to already include robustness evaluation at the stage of appropriate treatment technique and treatment plan selection.

The robustness tool assesses robustness according to multiple criteria. The dosimetric impact of uncertainties on target and OARs is investigated visually and quantitatively on a scenario-by-scenario base as well as in summary. The DVH viewer with the DVH bands including all investigated scenarios, the robustness maps and the RI contextualize the detailed evaluation according to dose-volume parameters, dose difference, and gamma passing rate of the individual scenarios. The gamma passing rate calculation is successfully validated against pymedphys and permits efficient assessment of the dosimetric impact of an uncertainty type. Differences in DVH parameter calculation, especially for small structures, arise due to the different volume calculations and the dose value: the triangular meshes are not constructed in the exact same way in the robustness tool and in Eclipse. Furthermore, Eclipse interpolates the dose between the voxels for DVH calculation and the robustness tool does not. In literature more complex interpolation methods, for example, the polynomial chaos expansion[39] or (quasi-) Monte Carlo methods,[40] as compared to our interpolation method are described, usually operating on the level of dose distributions. However, these methods assume that the probability distribution function of the uncertainties is known beforehand, and that the uncertainties are independent of each other. In general, the independence of the uncertainties cannot be guaranteed and there is limited knowledge, especially about the cumulative probability distribution of several uncertainties. The presented applications of the robustness tool confirm that restricting robustness investigations on the target volume[12] gives a misleading picture of the plan robustness, and simultaneous evaluation of dosimetric robustness

TABLE 5 Comparison of the DVH parameter evaluation: robustness tool (bold, left) and Eclipse (right)

	Planning target volume		Clinical target volume		Brain		Brainstem		Chiasma		Body	
Volume [cm^3]	**272.9**	272.4	**166.4**	165.5	**1479.5**	1479.1	**39.8**	39.3	**0.5**	0.1	**7599.6**	7508.2
*D*2% [Gy]	**63.3**	62.5	**62.2**	62.1	**61.8**	61.3	**28.7**	27.7	**27.3**	24.1	**59.8**	59.9
*D*98% [Gy]	**55.4**	56.6	**57.4**	58.2	**0.0**	1.3	**0.0**	0.4	**13.7**	14.6	**0.0**	0.1
D_{mean} [Gy]	**59.5**	59.9	**60.5**	60.1	**25.7**	26.2	**7.4**	6.4	**18.1**	16.8	**7.6**	8.0

on OARs is needed for a comprehensive robustness analysis of a treatment plan.

The application of the robustness tool fits into existing evaluation strategies: the flexibility in incorporating numerous uncertainty scenarios in the dose calculation and their evaluation makes it compatible with existing robustness approaches such as the "good practice scenario selection" and the "statistically sound scenario selection" evaluated by Sterpin et al.[41] However, the tool extends these strategies by considering new potential scenarios, various uncertainties, different treatment techniques, and flexible evaluation including the RI. The RI facilitates the robustness evaluation, summarizes, and efficiently compares the robustness of competing treatment plans and techniques for a given robustness space. Robustness evaluation is directly depending on the selected robustness space along with the conditions as robustness acceptance criteria. The RI is consequently also depending on those settings. Robustness evaluation by means of RI should always be reported with information about robustness space and the robustness conditions. Additionally, if the user of the robustness tool has knowledge about the probability distribution of the scenarios, a confidence interval for the RI can be reported.

We understand that the full potential of the RI only unfolds when standardized robustness spaces in terms of uncertainty axes, scenarios, and evaluation range, as well as relevant evaluation conditions such as $D98\%$, $D2\%$, or D_{mean} for the structures are defined. This standardizes facilitates and streamlines the comparison of RI of different robustness studies in various radiotherapy centers. Furthermore, the RI has the potential to serve as a threshold action level for replanning, discarding the plan, or recommending specific QA tests.

The robustness tool is expected to play a key role in our group in the field of robust optimization, the margin concept, and the implementation of new treatment techniques. With the help of this tool, we aim to gain an understanding of meaningful robust plan optimization by investigating treatment type-specific correlations between different uncertainties and by studying the sensitivity of a plan to different uncertainty types. Additionally, with the tool there is the potential to investigate the margin concept of target volumes and OARs toward flexible margins.[42] Finally, the robustness of new treatment techniques is explorable in order to facilitate their safe clinical implementation.

5 | CONCLUSION

In conclusion, by combining dose calculation of treatment plans using different treatment techniques, including patient and machine-related uncertainties, and flexible evaluation according to user-defined criteria, this robustness tool provides accurate comprehensive robustness evaluation, and fills the need for an overarching robustness evaluation tool to determine plan quality.

ACKNOWLEDGMENTS

This work was supported by Varian Medical Systems. The MC dose calculations were performed on UBELIX (www.id.unibe.ch/hpc), the high-performance computing cluster at the University of Bern.

Open access funding provided by Universitat Bern.

CONFLICT OF INTEREST

The authors have no relevant conflicts of interest to disclose.

REFERENCES

1. Hernandez V, Hansen CR, Widesott L, 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://doi.org/10.1016/j.radonc.2020.09.038
2. Schmidhalter D, Malthaner M, Born EJ, et al. Assessment of patient setup errors in IGRT in combination with a six degrees of freedom couch. *Z Med Phys*. 2014;24(2):112-122. https://doi.org/10.1016/j.zemedi.2013.11.002
3. Kerns JR, Childress N, Kry SF. A multi-institution evaluation of MLC log files and performance in IMRT delivery. *Radiat Oncol*. 2014;9(1):176. https://doi.org/10.1186/1748-717X-9-176
4. Yock AD, Mohan R, Flampouri S, et al. Robustness analysis for external beam radiation therapy treatment plans: describing uncertainty scenarios and reporting their dosimetric consequences. *Pract Radiat Oncol*. 2019;9(4):200-207. https://doi.org/10.1016/j.prro.2018.12.002
5. Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *Br J Radiol*. 2011;84(1007):967-996. https://doi.org/10.1259/bjr/22373346
6. Unkelbach J, Bortfeld T, Craft D, et al. Optimization approaches to volumetric modulated arc therapy planning. *Med Phys*. 2015;42(3):1367-1377. https://doi.org/10.1118/1.4908224
7. Pigorsch SU, Kampfer S, Oechsner M, et al. Report on planning comparison of VMAT, IMRT and helical tomotherapy for the ESCALOX-trial pre-study. *Radiat Oncol*. 2020;15(1):253. https://doi.org/10.1186/s13014-020-01693-2
8. Smyth G, Evans PM, Bamber JC, Bedford JL. Recent developments in non-coplanar radiotherapy. *Br J Radiol*. 2019;92(1097):20180908. https://doi.org/10.1259/bjr.20180908
9. Fix MK, Frei D, Volken W, et al. Part 1: Optimization and evaluation of dynamic trajectory radiotherapy. *Med Phys*. 2018;45(9):4201-4212. https://doi.org/10.1002/mp.13086
10. Mueller S, Fix MK, Joosten A, et al. Simultaneous optimization of photons and electrons for mixed beam radiotherapy. *Phys Med Biol*. 2017;62(14):5840-5860. https://doi.org/10.1088/1361-6560/aa70c5
11. Mueller S, Manser P, Volken W, et al. Part 2: Dynamic mixed beam radiotherapy (DYMBER): photon dynamic trajectories combined with modulated electron beams. *Med Phys*. 2018;45(9):4213-4226. https://doi.org/10.1002/mp.13085
12. van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol*. 2004;14(1):52-64. https://doi.org/10.1053/j.semradonc.2003.10.003
13. Wojtasik AM, Bolt M, Clark CH, Nisbet A, Chen T. Multivariate log file analysis for multi-leaf collimator failure prediction in radiotherapy delivery. *Phys Imaging Radiat Oncol*. 2020;15:72-76. https://doi.org/10.1016/j.phro.2020.07.011
14. Imae T, Haga A, Watanabe Y, et al. Retrospective dose reconstruction of prostate stereotactic body radiotherapy using

cone-beam CT and a log file during VMAT delivery with flattening-filter-free mode. *Radiol Phys Technol*. 2020;13(3):238-248. https://doi.org/10.1007/s12194-020-00574-3
15. Katsuta Y, Kadoya N, Fujita Y, et al. Log file-based patient dose calculations of double-arc VMAT for head-and-neck radiotherapy. *Phys Med*. 2018;48:6-10. https://doi.org/10.1016/j.ejmp.2018.03.006
16. Katsuta Y, Kadoya N, Fujita Y, et al. Patient-specific quality assurance using monte carlo dose calculation and elekta log files for prostate volumetric-modulated arc therapy. *Technol Cancer Res Treat*. 2017;16(6):1220-1225. https://doi.org/10.1177/1533034617745250
17. Pan Y, Yang R, Zhang S, et al. National survey of patient specific IMRT quality assurance in China. *Radiat Oncol*. 2019;14(1):69. https://doi.org/10.1186/s13014-019-1273-5
18. Katsuta Y, Kadoya N, Fujita Y, et al. Log file-based patient dose calculations of double-arc VMAT for head-and-neck radiotherapy. *Phys Med*. 2018;48:6-10. https://doi.org/10.1016/J.EJMP.2018.03.006
19. Imae T, Haga A, Watanabe Y, et al. Retrospective dose reconstruction of prostate stereotactic body radiotherapy using cone-beam CT and a log file during VMAT delivery with flattening-filter-free mode. *Radiol Phys Technol*. 2020;13(3):238-248. https://doi.org/10.1007/S12194-020-00574-3
20. LoSasso T, Chui CS, Ling CC. Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy. *Med Phys*. 1998;25(10):1919-1927. https://doi.org/10.1118/1.598381
21. Luo W, Li J, Price RA, et al. Monte Carlo based IMRT dose verification using MLC log files and R/V outputs. *Med Phys*. 2006;33(7Part1):2557-2564. https://doi.org/10.1118/1.2208916
22. Andreo P. Monte Carlo simulations in radiotherapy dosimetry. *Radiat Oncol*. 2018;13(1):121. https://doi.org/10.1186/s13014-018-1065-3
23. Fix MK, Manser P, Frei D, Volken W, Mini R, Born EJ. An efficient framework for photon Monte Carlo treatment planning. *Phys Med Biol*. 2007;52(19):N425. https://doi.org/10.1088/0031-9155/52/19/N01
24. Fix K, Manser P, Frei D, Volken W, Mini R, Born J. Efficient photon treatment planning by the use of Swiss Monte Carlo plan. *J Phys Conf Ser*. 2007;74(1):021004. https://doi.org/10.1088/1742-6596/74/1/021004
25. Giglioli FR, Garibaldi C, Blanck O, et al. Dosimetric multicenter planning comparison studies for stereotactic body radiation therapy: methodology and future perspectives. *Int J Radiat Oncol Biol Phys*. 2020;106(2):403-412. https://doi.org/10.1016/j.ijrobp.2019.10.041
26. van Herk M, Witte M, van der Geer J, Schneider C, Lebesque JV. Biologic and physical fractionation effects of random geometric errors. *Int J Radiat Oncol Biol Phys*. 2003;57(5):1460-1471. https://doi.org/10.1016/J.IJROBP.2003.08.026
27. Fredriksson A. A characterization of robust radiation therapy treatment planning methods—from expected value to worst case optimization. *Med Phys*. 2012;39(8):5169-5181. https://doi.org/10.1118/1.4737113
28. Magaddino V, Manser P, Frei D, et al. Validation of the Swiss Monte Carlo Plan for a static and dynamic 6 MV photon beam. *Z Med Phys*. 2011;21(2):124-134. https://doi.org/10.1016/j.zemedi.2010.10.010
29. Blanchette J. C++ GUI Programming with Qt 4. 2nd ed. 2008.
30. Ju T, Simpson T, Deasy JO, Low DA. Geometric interpretation of the γ dose distribution comparison technique: interpolation-free calculation. *Med Phys*. 2008;35(3):879-887. https://doi.org/10.1118/1.2836952
31. Manser P, Frauchiger D, Frei D, Volken W, Terribilini D, Fix MK. Dose calculation of dynamic trajectory radiotherapy using Monte Carlo. *Z Med Phys*. 2019;29(1):31-38. https://doi.org/10.1016/j.zemedi.2018.03.002
32. Wendling M, Zijp LJ, McDermott LN, et al. A fast algorithm for gamma evaluation in 3D. *Med Phys*. 2007;34(5):1647-1654. https://doi.org/10.1118/1.2721657
33. Jeraj R, Keall P. The effect of statistical uncertainty on inverse treatment planning based on Monte Carlo dose calculation. *Phys Med Biol*. 2000;45(12):3601. https://doi.org/10.1088/0031-9155/45/12/307
34. Pogson EM, Aruguman S, Hansen CR, et al. Multi-institutional comparison of simulated treatment delivery errors in ssIMRT, manually planned VMAT and autoplan-VMAT plans for nasopharyngeal radiotherapy. *Phys Med*. 2017;42:55-66. https://doi.org/10.1016/J.EJMP.2017.08.008
35. Jensen CA, Roa AMA, Johansen M, Lund JÅ, Frengen J. Robustness of VMAT and 3DCRT plans toward setup errors in radiation therapy of locally advanced left-sided breast cancer with DIBH. *Phys Med*. 2018;45:12-18. https://doi.org/10.1016/J.EJMP.2017.11.019
36. Qian J, Lee L, Liu W, et al. Dose reconstruction for volumetric modulated arc therapy (VMAT) using cone-beam CT and dynamic log files. *Phys Med Biol*. 2010;55(13):3597. https://doi.org/10.1088/0031-9155/55/13/002
37. Lee L, Le QT, Xing L. Retrospective IMRT dose reconstruction based on cone-beam CT and MLC log-file. *Int J Radiat Oncol Biol Phys*. 2008;70(2):634-644. https://doi.org/10.1016/J.IJROBP.2007.09.054
38. Katsuta Y, Kadoya N, Fujita Y, et al. Quantification of residual dose estimation error on log file-based patient dose calculation. *Phys Med*. 2016;32(5):701-705. https://doi.org/10.1016/J.EJMP.2016.04.015
39. Xiu D, Hesthaven JS. High-order collocation methods for differential equations with random inputs. *SIAM J Sci Comput*. 2006;27(3):1118-1139. https://doi.org/10.1137/040615201
40. Lemieux C. *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer; 2009. https://doi.org/10.1007/978-0-387-78165-5
41. Sterpin E, Rivas ST, van den Heuvel F, George B, Lee JA, Souris K. Development of robustness evaluation strategies for enabling statistically consistent reporting. *Phys Med Biol*. 2021;66(4):045002. https://doi.org/10.1088/1361-6560/ABD22F
42. Ahmad R, Bondar L, Voet P, et al. A margin-of-the-day online adaptive intensity-modulated radiotherapy strategy for cervical cancer provides superior treatment accuracy compared to clinically recommended margins: a dosimetric evaluation. *Acta Oncol (Madr)*. 2013;52(7):1430-1436. https://doi.org/10.3109/0284186X.2013.813640

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Loebner HA, Volken W, Mueller S, et al. Development of a Monte Carlo based robustness calculation and evaluation tool. *Med Phys*. 2022;49:4780–4793. https://doi.org/10.1002/mp.15683