ACTA ONCOLOGICA, 2017
VOL. 56, NO. 6, 769–778
http://dx.doi.org/10.1080/0284186X.2017.1290825

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Incorporation of relative biological effectiveness uncertainties into proton plan robustness evaluation

Jakob Ödén[a,b], Kjell Eriksson[b] and Iuliana Toma-Dasu[a,c]

[a]Department of Physics, Medical Radiation Physics, Stockholm University, Stockholm, Sweden; [b]RaySearch Laboratories, Stockholm, Sweden; [c]Department of Oncology and Pathology, Medical Radiation Physics, Karolinska Institutet, Stockholm, Sweden

ABSTRACT

Background: The constant relative biological effectiveness (RBE) of 1.1 is typically assumed in proton therapy. This study presents a method of incorporating the variable RBE and its uncertainties into the proton plan robustness evaluation.

Material and methods: The robustness evaluation was split into two parts. In part one, the worst-case physical dose was estimated using setup and range errors, including the fractionation dependence. The results were fed into part two, in which the worst-case RBE-weighted doses were estimated using a Monte Carlo method for sampling the input parameters of the chosen RBE model. The method was applied to three prostate, breast and head and neck (H&N) plans for several fractionation schedules using two RBE models. The uncertainties in the model parameters, linear energy transfer and α/β were included. The resulting DVH error bands were compared with the use of a constant RBE without uncertainties.

Results: All plans were evaluated as robust using the constant RBE. Applying the proposed methodology using the variable RBE models broadens the DVH error bands for all structures studied. The uncertainty in α/β was the dominant factor. The variable RBE also shifted the nominal DVHs towards higher doses for most OARs, whereas the direction of this shift for the clinical target volumes (CTVs) depended on the treatment site, RBE model and fractionation schedule. The average RBE within the CTV, using one of the RBE models and 2 Gy(RBE) per fraction, varied between 1.11–1.26, 1.06–1.16 and 1.14–1.25 for the breast, H&N and prostate patients, respectively.

Conclusions: A method of incorporating RBE uncertainties into the robustness evaluation has been proposed. By disregarding the variable RBE and its uncertainties, the variation in the RBE-weighted CTV and OAR doses may be underestimated. This could be an essential factor to take into account, especially in normal tissue complication probabilities based comparisons between proton and photon plans.

ARTICLE HISTORY
Received 4 January 2017
Accepted 26 January 2017

Introduction

The finite range and sharp distal fall-off of proton beams may be exploited in radiotherapy planning to create conformal dose distributions with limited dose burden to organs at risk (OAR). However, these advantageous physical properties of proton beams also make them sensitive to uncertainties, such as patient misalignments and errors in the conversion from CT-numbers to proton stopping-power. This is of particular importance for intensity-modulated proton therapy (IMPT) due to potential negative interplay effects if several inhomogeneous dose distributions are misaligned. This effect may be mitigated by careful choices of the number of beams, beam angles, starting conditions in the spot optimization, image guidance procedure, CT-calibration method, etc. or by incorporating the setup and range uncertainties directly in a robust optimization approach [1]. A plan is then considered to be robust if the dose agreement between the nominal scenario (no errors) and appropriate error scenarios satisfies the dosimetric requirements for the clinical target volume (CTV) and the OARs. These worst-case dose distributions are commonly evaluated under a finite number of error scenarios of rigid setup and fixed range errors. However, since most treatment courses are fractionated, it can be expected that the uncertainty in the dose delivered will be reduced compared to a single fraction. This is due to the convergence of the random setup error and the negligible systematic setup error under the use of proper image guidance. This has recently been proposed to be incorporated in the robustness evaluation method [1].

Regardless of whether a robust optimization or careful iterative planning process is used to mitigate the sensitivity to uncertainties, the constant relative biological effectiveness (RBE) of 1.1 compared to photon therapy is typically used. However, the RBE depends on the chosen biological endpoint and dose per fraction, and there is increasing evidence that it also varies with the linear energy transfer (LET) and the tissue

CONTACT Jakob Ödén jakob.oden@fysik.su.se Department of Physics, Medical Radiation Physics, Stockholm University, Karolinska University Hospital P9:02, SE-171 76 Stockholm, Sweden

Supplemental data for this article can be accessed here.

type [2]. Several studies have reported substantial differences between the constant RBE of 1.1 and various variable RBE models for clinical cases [3–8]. Most of them have used models based on the linear-quadratic (LQ) survival curve. These models use parameters fitted to *in vitro* cell survival data and are functions of the physical proton dose per fraction, the dose-average LET (LET_d) and the tissue characteristic parameter α/β. However, few of the studies have incorporated the uncertainties in the models. Some have evaluated the impact of α/β variations by performing calculations using several values [6–9], but to our knowledge, no one has fully included the uncertainties in the physical proton dose, α/β, LET_d and model parameters into the robustness evaluation of the final RBE-weighted dose. Such a comprehensive evaluation could be an essential factor to include in the proton plan evaluation, especially in comparisons with photon plans. This is of special concern if the selection of patients for proton therapy is based on the reduction of side effects, as proposed by Langendijk et al. [10], since most models predict higher RBE values for low α/β tissues, which are associated with late toxicity [2].

This study proposes a method to incorporate the variable RBE and its uncertainties into proton plan robustness evaluation by splitting the evaluation process into two parts, a physical and a radiobiological part. The physical dose evaluation is essentially identical to the traditional evaluation, without the 1.1 factor included. The radiobiological step evaluates the RBE-weighted dose, obtained with the RBE model of choice, and its uncertainties using a Monte Carlo (MC) method. In this study, the constant RBE of 1.1 is accompanied by the LQ-based variable RBE models by Wedenberg et al. [11] and McNamara et al. [12] for the evaluation of the method.

Material and methods

The robustness evaluation was divided into two parts. In the first part, the worst-case scenario of the physical proton dose was evaluated per voxel in the presence of setup and range uncertainty, including the fractionation dependence. In the second step, the uncertainty in the RBE-weighted dose was evaluated per voxel by taking the RBE uncertainties into account using a MC method.

Robustness of the physical proton dose

The error bound of the physical dose distributions due to setup uncertainties was calculated by employing 14 rigid patient shifts, and including the effect of the number of fractions. The shifts were made in the positive and negative directions along the principal axes (six shifts) and along the diagonal axes for all triplets possible (eight shifts) [1]. The combined standard deviation (SD) of the positioning error used in this study was 2 mm for head and neck (H&N) patients and 3.5 mm for prostate and breast patients. The systematic setup error was neglected, which is considered reasonable when using a daily setup correction protocol. Employing the findings by Lowe et al. [1], resulted in shifts along the principal axes of 1.1 or 2.0 mm (0.56σ) and 2.9 or 5.0 mm (1.44σ) along the diagonal axes in this study to ensure a dose error-space bounded by the commonly clinically used 85% confidence interval (CI) [13].

The 14 error scenarios resulted in a probability distribution of the dose for each voxel for a single fraction. For a fractionated treatment, this distribution was approximated with a normal distribution with a SD of σ_n equal to $\sigma_1/\sqrt{n}$, where σ_1 is the SD for the single fraction distribution and n is the number of fractions. The worst-case lower and upper physical dose in a voxel was then estimated to be the mean value $\pm 3.3\sigma_n$, constrained by the single fraction dose error as maximum and minimum, and always encompassing the nominal dose value for the voxel. This corresponds to the 99.9% CI and ensured that the overall limit was determined by the error magnitude chosen for the initial error scenarios (i.e., 85% CI here). The method is described in detail in the original work [1].

In this study, the robustness to setup uncertainties was evaluated separately for three range scenarios, 0 and ±3% error in the nominal densities in the patients' CT data. All three scenarios resulted in one lower and one upper worst-case physical dose for each voxel, which were saved for use in the radiobiological evaluation step. This in contrast to linearly adding the uncertainty resulting from systematic range errors [1].

All error scenarios were calculated using a research version of the treatment planning system RayStation v4.6 (RaySearch Laboratories, Stockholm, Sweden). The doses were exported via the use of the IronPython scripting tool available inside RayStation to MATLAB (The Mathworks, Inc., Natwick, MA), where the remaining steps were implemented.

Robustness of the RBE-weighted proton dose

To incorporate the variable RBE and its uncertainties into the error bound for each voxel, a MC method is proposed. In this study, the MC robustness method was applied using the variable RBE models by Wedenberg et al. [11] and McNamara et al. [12], which both are based on LQ survival curves of cell lines irradiated *in vitro*. The Wedenberg model consists of one model parameter (q) and has no variation with LET_d or α/β of the quadratic dose term, whereas the McNamara model uses four parameters (p_0, p_1, p_2 and p_3) and has a LET_d and α/β dependence of the quadratic dose term. Their equations and parameter values are stated in the Supplementary data. The assumed nominal values of α/β are stated in Supplementary Table S1 together with the estimated 95% CIs, taking the known variability into account [14–22]. The LET_d was calculated for the three range error scenarios (no setup error) for each voxel inside the research version of RayStation using an already existing MC dose framework [8].

To perform the MC robustness method, probability density functions (PDF) have to be generated for the parameters included in the RBE model of choice. In this study, PDFs were generated for the model parameters using the data in the Supplementary material, for α/β using the nominal values and 95% CIs from Supplementary Table S1, for the LET_d assuming a 95% CI of ±10% of the nominal LET_d for each voxel and range scenario evaluated (0 or ±3%), which was derived from

an analysis of MC calculations of several patient shifts. The parameters with a symmetric CI (LET_d per voxel, p_i and α/β for most of the tissues) were approximated with a normally distributed PDF using the nominal value as mean value and the 95% CI to determine the SD. To ensure positive LET_d and α/β values, all such normal PDFs were truncated bilaterally at the distance from the mean value to zero. The asymmetrical distributions (q and α/β for prostate CTV) were approximated with a log-normal PDF, fitted to fulfill the 2.5 percentiles given by the estimated 95% CI, with ensured positive values.

After generating the PDFs, repeated simulations were performed using pseudo-random values drawn from each PDF. We performed 10,000 simulations for each of the six physical dose boundaries for a voxel (one lower and one upper for each range error scenario). Each of these simulations resulted in a RBE value per voxel, given from the RBE model of choice, which subsequently resulted in a RBE-weighted dose value per voxel when multiplied with the worst-case physical dose in the corresponding voxel. Each of the six scenarios thereby resulted in 10,000 RBE-weighted dose values per voxel. Depending on the RBE model and PDFs used, each of the six distributions of RBE-weighted dose values per voxel could be truncated at the desired confidence level to avoid unrealistic RBE values originated from combinations of input parameters sampled far from the mean values. In this study, we used a 95% voxel-wise CI around the mean RBE-weighted dose by truncating the distributions at the lower and upper 2.5 percentiles. The minimum and maximum values of the remaining 57,000 values per voxel were then stored as the lower and upper boundaries of the RBE-weighted dose for each voxel, from which cumulative lower and upper DVHs were generated for all ROIs. The implementation of this MC method was done in MATLAB using the embedded method 'random' to generate pseudo-random numbers from PDFs generated using the method 'makedist'.

The robustness using the constant RBE of 1.1 was calculated without any uncertainties, hence by only upscaling the worst-case physical dose from step one with 10% for each voxel for all the six physical dose boundaries. The maximum and minimum of these six values were stored as the lower and upper boundaries for each voxel. From these boundaries, cumulative lower and upper DVHs and normal tissue complication probabilities (NTCP) ranges were generated for all region of interests (ROIs). Ranges of NTCP were also calculated using the dose boundaries for the lung, parotid glands and rectum, for both the variable RBE models and the constant RBE. The NTCP parameters used in this study are found in Supplementary Table S2 [23–25].

Patients and treatment plans

The robustness method was applied to three breast, three bilateral H&N and three prostate plans. The CTV included the whole left breast for the breast patients, the entire prostate gland for the prostate patients and the delineated low- and high-risk volumes of the H&N patients, including bilateral neck nodes. The planning target volume (PTV) was defined as a uniform expansion of the CTV with 5 mm, except in the beam directions for the prostate patients where a margin corresponding to 3%+1 mm was used. The PTV margin of 5 mm in the H&N and breast cases approximately corresponds to 3%+1 mm as the ranges were ~10 cm. Additionally both lungs, contralateral breast, heart and left anterior descending (LAD) artery were delineated as OARs for the breast patients. For the prostate patients, rectum, bladder, femoral heads and penile bulb were delineated, and spinal cord, brainstem, parotid glands, submandibular glands, oral cavity, larynx and mandible were defined for the H&N patients. All volumes were delineated by experienced oncologists on the patient's CT image data set, which had 2–3 mm slice thickness.

Proton plans, using pencil beam scanning (PBS), were generated for the nine patients in the research version of RayStation. The beam data used was obtained from a PBS dedicated nozzle manufactured by IBA (Ion Beam Applications S.A., Louvain-la-Neuve, Belgium) with proton energies between 70 and 230 MeV available. The breast plans consisted of three oblique IMPT fields (~340, 20 and 60°), the H&N plans of four IMPT fields (~70, 120, 240 and 290°) and the prostate plans of two opposing IMPT fields (90 and 270°). To ensure superficial target coverage, a range shifter of 4 cm water equivalent thickness was used for all breast fields and the two anterior H&N fields. The prescribed dose (PD) was normalized to the median dose ($D_{50\%}$) of the PTV for all plans. The PTV coverage goals used were $D_{98\%} \geq 95\%$ of the PD and $D_{2\%} \leq 105\%$ of the PD. The clinical goals for the OARs were generally adopted from the QUANTEC summary, but the doses were always tried to be kept to a minimum after ensuring fulfillment of the PTV goals.

The physical proton dose was optimized using the clinically used pencil beam algorithm under the assumption of a constant RBE of 1.1. A fractionation dose of 2 Gy(RBE) in 25, 35 and 39 fractions were prescribed to the CTV of the breast, H&N and prostate, respectively, whereas a fractionation dose of 1.6 Gy(RBE) was prescribed to the H&N low-risk CTV.

Fractionation dependence

To evaluate the impact of the fractionation schedule, the plans were renormalized for schedules between 5–25 fractions for the breast cases, 35–70 fractions for the H&N cases and 5–39 fractions for the prostate cases. The dose per fraction for the breast and prostate cases was calculated to keep the equivalent total dose delivered in 2 Gy(RBE) fractions ($EQD2_{\alpha/\beta}$) equal to the reference schedule (without repopulation). An α/β of 4 Gy was used for the breast patients, whereas 5 Gy was used for the prostate patients, in order to keep the predicted late lung and rectal toxicity approximately constant. The $EQD2_{\alpha/\beta}$ is calculated using Equation (1),

$$EQD2_{\alpha/\beta} = \frac{n \cdot d\left(1+\frac{d}{\alpha/\beta}\right) - \frac{\ln 2}{\alpha}\frac{(T-T_k)}{T_d}}{\left(1+\frac{2}{\alpha/\beta}\right)}, \quad (1)$$

where, n is the number of fractions, d is the RBE-weighted dose per fraction (RBE = 1.1), α and β are the linear and dose-squared coefficients of dose, T is the overall time, T_k is the kick-off time for accelerated repopulation and T_d is the average doubling time. If no repopulation was assumed, or if $T < T_k$, the second term in the numerator was set to zero.

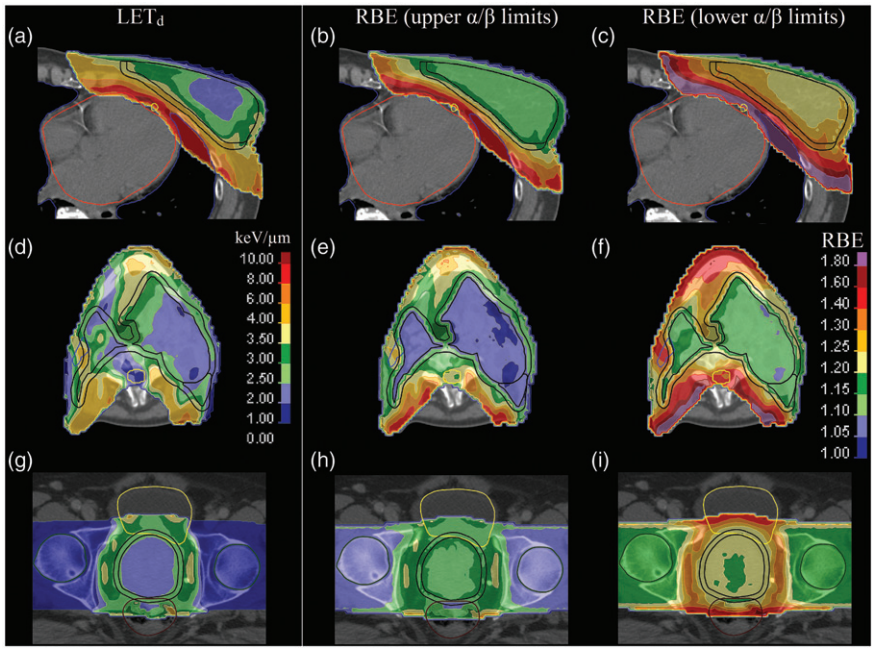

Figure 1. Representative transverse slices for one patient of each treatment site showing the nominal LET_d distribution (left column), the resulting RBE distributions using the Wedenberg model with the upper limits of the 95% CIs of α/β (middle column) and the resulting RBE distributions using the Wedenberg model with the lower limits of the 95% CIs of α/β (right column). The PD was 2 Gy(RBE) per fraction and only voxels receiving $\geq$5% of the PD are displayed.

In the H&N cases, the dose per fraction was calculated not to exceed the $EQD2_{3Gy}$ (without repopulation) or the $EQD2_{10Gy}$ (with repopulation) of the reference schedule of 70 Gy in 35 fractions delivered in six fractions per week over 39 days. This approach was used to keep the general late toxicity and the acute oral mucosal toxicity on the same, or lower, level compared to the reference schedule. The parameters assumed for human mucosa were $\alpha/\beta = 10$ Gy, $\alpha = 0.35\,Gy^{-1}$, $T_k = 7$ days and $T_p = 2.5$ days [26]. Six fractions per week were assumed for 42 or less fractions, overall time of 46 days for 43–59 fractions and 10 fractions per week (BID) for 60 or more fractions.

Results

Implementation of the MC method

The MC method was successfully implemented in MATLAB. No discrepancies were found when comparing MATLAB generated RBE-weighted DVHs with the corresponding DVHs calculated using RayStation and an IronPython script. The distributions of the pseudo-random generated parameters were plotted and evaluated for each simulation to assure correct sampling. Supplementary Figure S1 shows examples of pseudo-random generated histograms together with the corresponding theoretical PDFs for α/β, LET_d and the q-parameter in the Wedenberg model and the resulting histogram, using the MC method, of the RBE-weighted dose for a voxel in the prostate CTV.

Plan evaluation and α/β dependence

All proton plans fulfilled the clinical goals for the PTV assuming a constant RBE of 1.1, with satisfying OAR doses. The average nominal RBE within the CTVs was 1.16, 1.09 and 1.20 using the Wedenberg model and 1.17, 1.09 and 1.22 using the McNamara model for the breast, H&N and prostate patients, respectively. Figure 1 shows the LET_d distribution and the corresponding RBE distributions for the reference schedules of 2 Gy(RBE) per fraction for one patient of each treatment site, assuming the Wedenberg model and the upper and lower limits of the 95% CIs of the α/β for all tissues (from Supplementary Table S1). These distributions graphically illustrate the dominant effect of the α/β variation on the predicted RBE. When only including the α/β variation of the Wedenberg model in the simulations, the average RBE within the CTVs varied between 1.13–1.24, 1.07–1.14 and 1.16–1.23 for the breast, H&N and prostate patients, respectively. This could be compared with the corresponding total RBE variation, when including LET_d and model parameter uncertainty, of 1.11–1.26, 1.06–1.16 and 1.14–1.25. This α/β dominance was similar for simulations using the McNamara model.

Figure 2 shows the corresponding dose distributions obtained assuming RBE = 1.1 and the RBE values given by the Wedenberg model in Figure 1. The increased LET_d outside the CTVs seen in Figure 1(a,d,g) results in rims of high doses outside the CTV borders for all treatment sites when applying the variable RBE model (Figure 2). This effect is especially pronounced when the lower 95% CI limit of α/β (1–2 Gy) is assumed for the normal tissues. For the breast and prostate cases, substantially higher doses are also predicted inside the CTV when assuming the lower 95% CI limit of α/β (Figure 2(c,i)). In the H&N cases, the CTV doses are quite similar to the ones predicted by the constant RBE of 1.1 due to the relative high α/β associated with it (Figure 2(e,f)).

Robustness evaluation

All plans were evaluated as robust under the assumption of the constant RBE of 1.1 without uncertainties, both in terms of CTV coverage and OAR doses. The DVH error bands for the

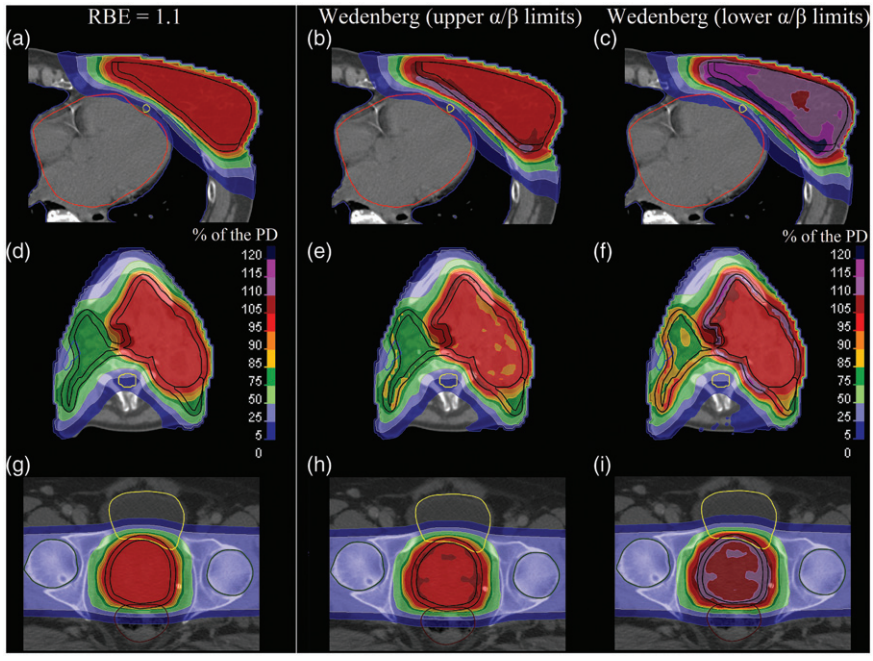

Figure 2. Representative transverse slices for one patient of each treatment site showing the nominal dose distribution using RBE =1.1 (left column), the Wedenberg model with the upper limits of the 95% CIs of α/β (middle column) and the Wedenberg model with the lower limits of the 95% CIs α/β (right column). The PD was 2 Gy(RBE) per fraction.

CTVs and one OAR per site, resulting from the worst-case dose distributions using the MC method with the two variable RBE models, are shown in Figure 3 for one patient of each treatment site. It could be observed that the two variable RBE models demonstrated consistent trends for all CTVs and OARs in terms of nominal DVHs and error bands, despite the slightly different model configurations. However, a minor shift in the nominal DVH and the error band is seen for the prostate CTV in Figure 3(e), where the McNamara model predicts a slightly higher RBE. Figure 4 shows the DVH error bands for two OARs per treatment site for one patient using the constant RBE of 1.1 (without uncertainty) and the Wedenberg model (including model uncertainties). As seen, when applying the MC method with the RBE uncertainties included, the DVH error bands broaden for all structures studied compared to using RBE = 1.1 without uncertainties. A selection of dosimetric data from the robustness evaluation is stated in Table 1 and Supplementary Table S3 for one patient of each treatment site. The trends of the nominal DVHs and error bands in Figures 3 and 4 are reflected with shifted nominal values and wider ranges in the dosimetric data for the variable RBE models compared with the constant RBE. These trends are also illustrated in Table 2, where the NTCP calculations are presented for all nine patients. However, the effect on the lung NTCPs are limited due to the very low doses given to the lung in the breast treatments. For the parotid glands, there is a substantial increase in the calculated NTCPs and their ranges when applying the variable RBE models, whereas the absolute effect is slightly smaller for the rectum.

Fractionation dependence

To facilitate the presentation of the results, two fractionation schedules per treatment site are presented together with the reference schedules of 2 Gy(RBE) per fraction. The schedules were chosen to keep the $EQD2_{\alpha/\beta}$ constant, according to the conditions described in the 'Material and Methods' section, and to cover the whole interval of fractionation schedules. Therefore, fractionation doses of 2.90 and 6.00 Gy(RBE) in 15 and 5 fractions where chosen for the breast patients, 1.75 and 1.15 Gy(RBE) in 42 and 70 fractions for the H&N patients and 3.29 and 8.24 Gy(RBE) in 20 and 5 fractions for the prostate patients.

Figure 5 shows the nominal DVHs and the corresponding error bands from the robustness analysis for the CTV and one OAR for one patient per treatment site (not the same patients as in Figures 3 and 4). The constant RBE of 1.1 is accompanied by the Wedenberg model. The calculated nominal NTCPs and their ranges are displayed in Table 3 for the constant RBE and both variable RBE models. The predicted variable RBE and its error bound generally decreases with increasing fractionation dose for both models. For high fractionation doses (e.g., 6 and 8.24 Gy) the Wedenberg model predicts even lower RBEs than 1.1 for the targets (Figure 5(c,i)), whereas the McNamara model predicts RBEs around 1.1 in those situations. This is also indicated by the rectal NTCP for the schedule of five fractions in Table 3, where the Wedenberg model predicts lower toxicity than using RBE = 1.1, whereas the McNamara model returns a value close to the constant 1.1 assumption.

Discussion

A novel method of incorporating RBE uncertainties into the proton plan robustness evaluation has been presented. Implementing this method generates estimates of the robustness of the RBE-weighted dose including physical and radiobiological uncertainties. Furthermore, the method is flexible

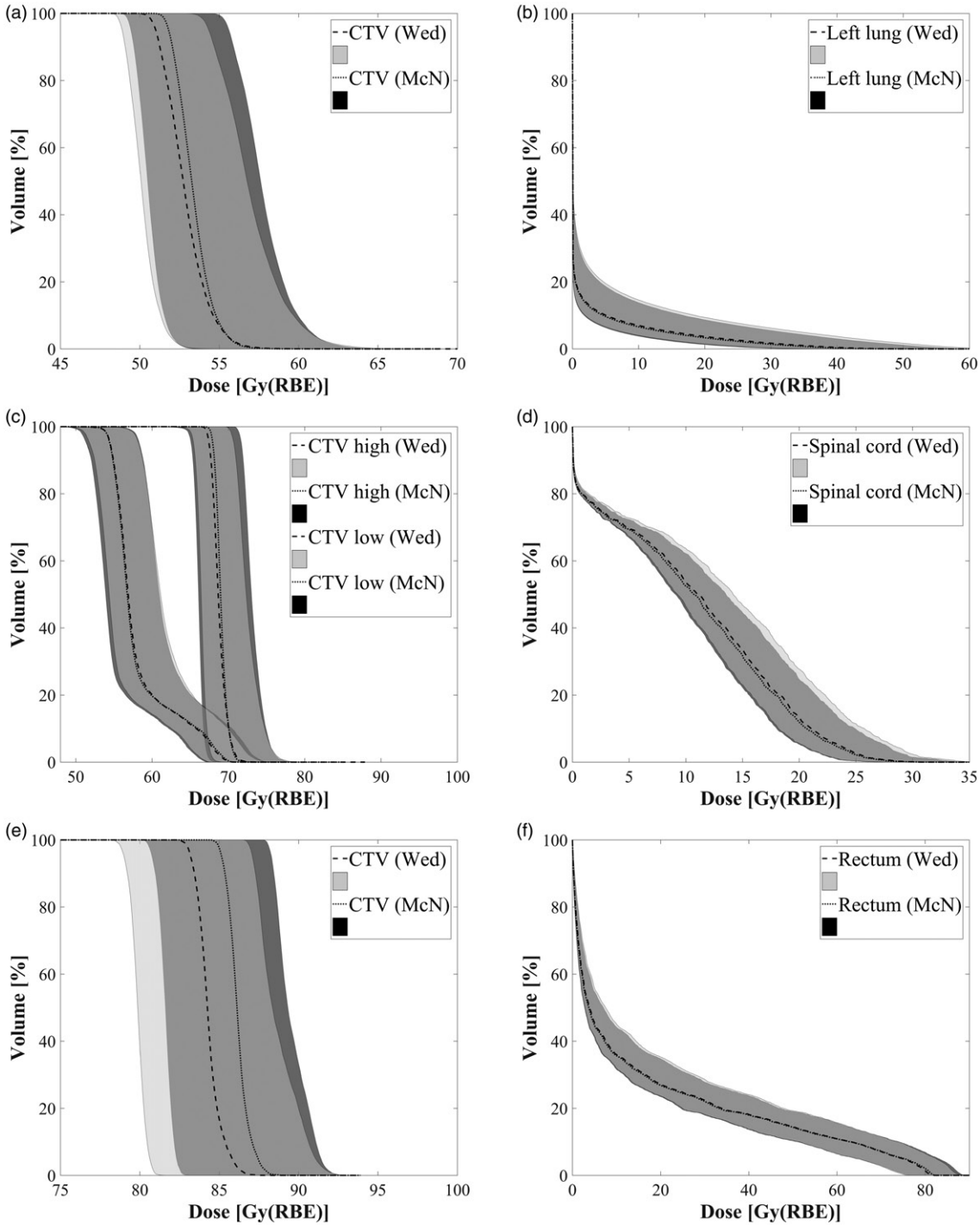

Figure 3. DVHs for the CTVs and one OAR for one patient of each treatment site using the Wedenberg and McNamara models. The left column shows the evaluation of the CTVs and the right column the OARs. The nominal DVHs for the Wedenberg (dashed lines) and for the McNamara model (dotted lines) are shown together with their corresponding error bands from the robustness evaluation.

and easy to adjust according to the settings desired by the user in terms of the programing language, RBE model, error scenarios, fractionation dependence and radiobiological uncertainties. Moreover, the method could easily be modified to include pseudo-random sampling of the physical dose per voxel, as for the other parameters, to generate the final voxel-wise RBE-weighted dose histogram. With an appropriate truncation at a representative percentile, these results are similar to the results using the method described here, where the worst-case physical doses are used. The method is presented in this format to separate the physical dose evaluation from the radiobiological evaluation and to clearly distinguish the addition made in this study to the traditional robustness evaluation.

Employing the method using two RBE models on three different treatment sites, with clinically realistic error scenarios and reasonable radiobiological uncertainties, indicates that the traditional robustness evaluation may underestimate the variation in the RBE-weighted doses for both the CTVs and the OARs (Figures 3–5), as well as in the predicted

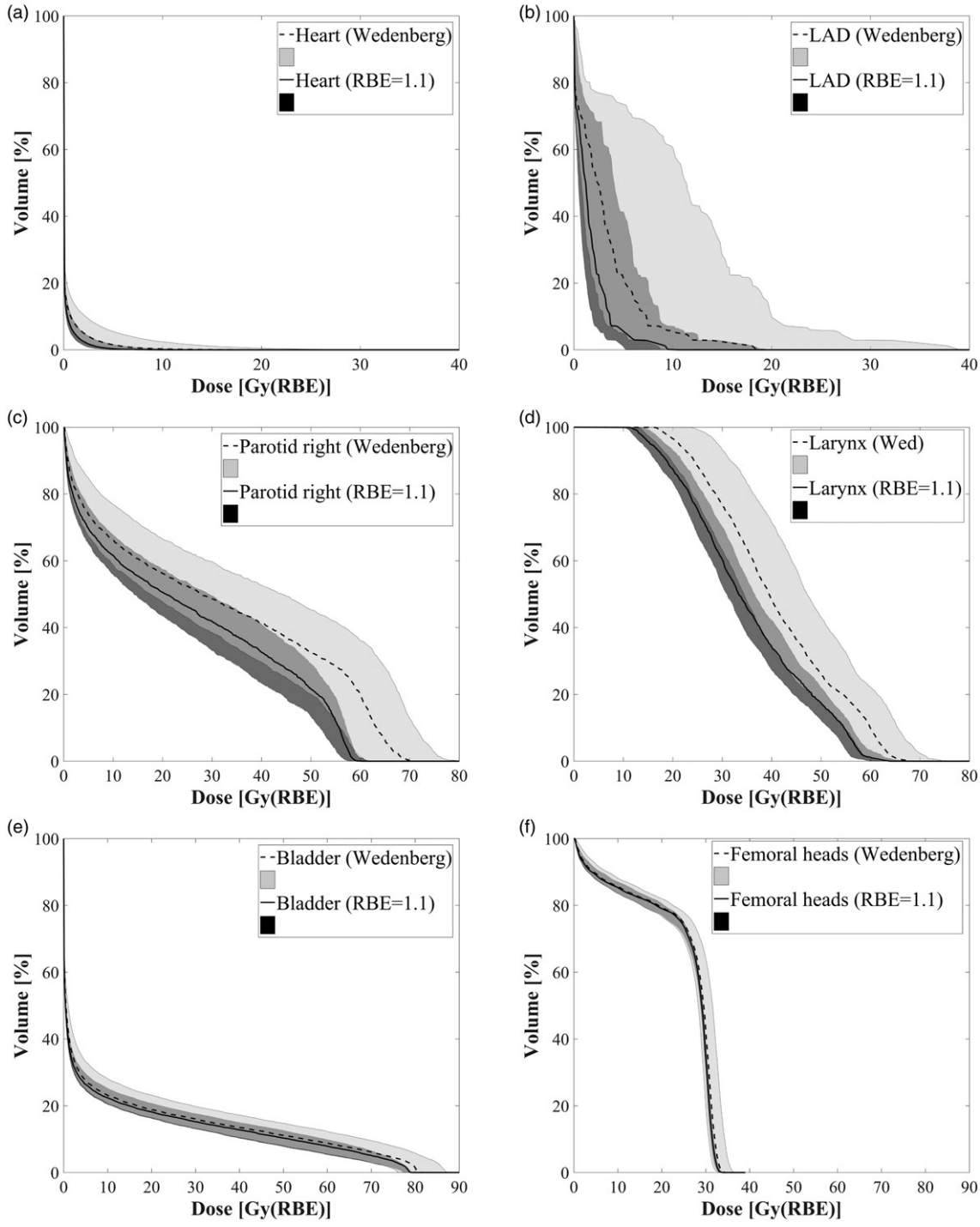

Figure 4. DVHs for two OARs for one patient per treatment site using the constant RBE and the Wedenberg model. Panels (a) and (b) are for the breast case, (c) and (d) for the H&N and (e) and (f) for the prostate case. The nominal DVHs for the constant RBE (solid lines) and for the Wedenberg model (dashed lines) are shown together with their corresponding error bands from the robustness evaluation.

NTCPs (Tables 2 and 3). The two models show similar trends and ranges and the dominant factor for both models is the uncertainty in the α/β, as indicated by results in previous studies [8,12] and Figures 1 and 2 in this study. This was also validated by additional simulations where only one parameter was allowed to vary at a time. The uncertainties in the fitted model parameters and the LET_d contribute minor to the resulting error bands, but should not be neglected in the analysis. However, as there exist extensive variability of the α/β values [14–22] and it turns out to be the dominant factor, the estimated PDFs are essential for this analysis. Therefore, we aimed to use realistic PDFs of the α/β values for each ROI in order to predict realistic error bounds in the final RBE-weighted doses and calculated NTCPs. However, when having a specific NTCP as a biological endpoint, the RBE model might also be different from the ones used in this study [2], which primarily are aimed to predict cell survival and hence TCP [11,12].

In addition to the two RBE models used in this study, there is a number of other models, as pointed out in the

Table 1. Nominal values and ranges of the mean dose from the robustness evaluation for the 2 Gy(RBE) fractionation schedule for one breast, one H&N and one prostate patient.

Patient	ROI	Dose metric	RBE 1.1	Wedenberg	McNamara
Breast	CTV	D_{mean} [Gy(RBE)]	50.0 (49.5–50.4)	52.9 (50.2–57.0)	53.4 (50.6–57.9)
	Heart	D_{mean} [Gy(RBE)]	0.18 (0.11–0.35)	0.32 (0.18–0.87)	0.30 (0.17–0.77)
	LAD	D_{mean} [Gy(RBE)]	1.66 (0.80–4.63)	3.28 (1.29–11.5)	2.98 (1.20–10.3)
	Left lung	D_{mean} [Gy(RBE)]	1.25 (0.78–2.56)	2.03 (1.11–5.07)	1.89 (1.05–4.60)
H&N	CTV high	D_{mean} [Gy(RBE)]	70.1 (69.3–70.8)	68.8 (66.5–72.6)	69.1 (66.1–73.3)
	CTV low	D_{mean} [Gy(RBE)]	58.7 (57.5–59.6)	58.3 (55.6–62.4)	58.2 (55.1–62.5)
	Larynx	D_{mean} [Gy(RBE)]	35.1 (32.5–38.1)	40.8 (35.6–48.3)	40.2 (35.3–47.3)
	Parotid	D_{mean} [Gy(RBE)]	25.0 (20.7–29.2)	30.7 (23.7–39.7)	30.0 (23.3–38.4)
	Spinal cord	D_{mean} [Gy(RBE)]	9.02 (8.28–9.82)	10.7 (9.11–13.4)	10.4 (9.06–12.9)
Prostate	CTV	D_{mean} [Gy(RBE)]	77.9 (77.3–78.6)	84.3 (79.8–88.8)	86.2 (81.7–89.6)
	Bladder	D_{mean} [Gy(RBE)]	11.0 (9.31–12.8)	11.8 (9.49–15.3)	11.7 (9.43–15.1)
	Femoral heads	D_{mean} [Gy(RBE)]	24.5 (23.8–25.2)	25.0 (23.4–27.1)	25.1 (23.5–27.4)
	Rectum	D_{mean} [Gy(RBE)]	15.8 (13.4–18.3)	17.0 (13.6–22.4)	16.9 (13.5–22.0)

Table 2. Nominal values and ranges of the NTCPs from the robustness evaluation for the nine patients for the fractionation schedules of 2 Gy(RBE).

		NTCP (%)		
ROI and endpoint	RBE model	Patient 1	Patient 2	Patient 3
Lung	RBE 1.1	0.38 (0.36–0.42)	0.39 (0.37–0.45)	0.38 (0.36–0.43)
Grade ≥2 pneumonitis	Wedenberg	0.40 (0.37–0.53)	0.43 (0.39–0.61)	0.41 (0.38–0.56)
	McNamara	0.40 (0.37–0.51)	0.42 (0.38–0.58)	0.40 (0.37–0.53)
		Patient 4	Patient 5	Patient 6
Parotid gland (best spared)	RBE 1.1	17.6 (11.5–25.1)	10.5 (8.3–13.4)	27.6 (21.4–33.9)
	Wedenberg	28.2 (15.5–49.6)	15.7 (10.4–26.4)	38.4 (26.0–56.1)
Flow rate <25%	McNamara	26.7 (14.9–46.3)	15.0 (10.1–24.5)	37.4 (25.6–53.9)
		Patient 7	Patient 8	Patient 9
Rectum	RBE 1.1	2.9 (0.8–3.7)	3.3 (1.1–4.3)	1.7 (0.6–2.7)
Grade ≥2 rectal toxicity	Wedenberg	4.7 (0.7–17.7)	5.3 (0.9–20.1)	2.3 (0.5–10.9)
	McNamara	5.2 (0.8–19.1)	6.0 (1.0–21.6)	2.8 (0.5–12.1)

review by Paganetti [2]. Even though it is beyond the scope of this study to present a comparison between all models, it is worth pointing out that some important trends such as increased RBE with lower α/β, lower dose and higher LET are common for most models. The magnitude is, however, model dependent and hence could the resulting DVH error bands using our method depend on the chosen RBE model.

The radiobiological uncertainties used in this study could potentially be incorporated directly into a robust optimization. Doing so, one would have the opportunity to optimize the physical dose so that the resulting RBE-weighted dose would be robust against the worst-case scenarios. If this could be combined with a LET optimization, it could be even further explored and the predicted RBEs for the OARs could potentially be lowered, as they often have the highest LET_d values (Figure 1). This would lower the nominal RBE for the OARs, but also result in narrower error bands as the α/β dependence of the models is reduced with decreasing LET_d values [11,12]. This is indicated in Figure 5, where e.g., the error bound for the femoral heads is much narrower than for parotid gland, larynx and LAD, in spite of the fact that they all have the same assumed PDF of the α/β. This is explained by the low average LET_d of 1.30 keV/μm in the femoral head, compared to 3.35, 3.72 and 7.47 keV/μm in the parotid gland, larynx and LAD, respectively. If the LET_d could be lowered in these ROIs, the resulting error bound would be narrower even with the same assumption of α/β values.

The magnitude of the error bounds could be even further explored by using alternative fractionation schedules. As the RBE decreases with increasing dose, it also becomes less dependent on the α/β value for higher fractionation doses [5,8,11,12]. On the other hand, the sensitivity to random setup errors becomes more important as the number of fractions decreases [1]. These two effects compete with each other, as seen in Table 3, where the range of the lung and rectal NTCPs widens for the constant RBE as the number of fractions decreases due to the fractionation effect. On the contrary, the ranges becomes narrower for the two variable RBE models since the decrease in RBE with increasing dose is more pronounced than the increase due to the fractionation effect. This could be further explored if implemented in a spatiotemporal optimization [27].

Although the uncertainties in the RBE models and NTCP calculations are substantial, this study indicates that the variable RBE models shift the nominal dose towards higher doses and increase the error bounds for all OARs studied. As it is likely that the clinical benefit of a proton treatment often will be compared with a substantially cheaper photon treatment, these effects could be of importance in a NTCP-based selection system [10]. In order to perform a more comprehensive selection, the uncertainties in the NTCP predictions should preferably be included in such an analysis.

Ultimately, if introducing a variable RBE model in the clinic, the physical doses required to obtain a satisfying RBE-weighted dose might be different from using the constant RBE of 1.1, as the predicted RBE may differ from 1.1. This is a delicate clinically issue and would require a correct and well-

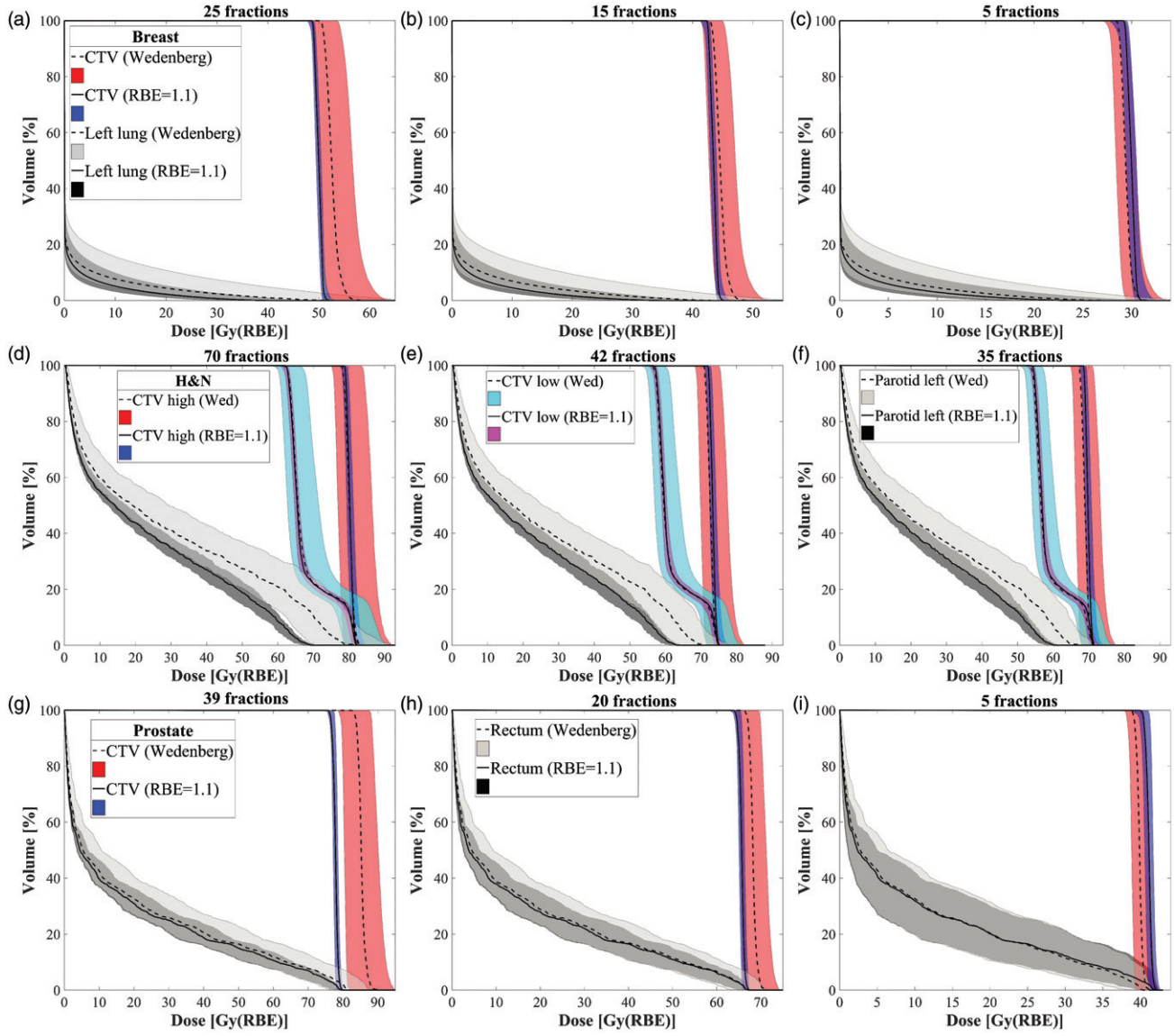

Figure 5. DVHs for the CTVs and one OAR for three fractionation schedules for one patient of each treatment site. Panels (a)–(c) is for the breast patient, (d)–(f) for the H&N patient and (g)–(i) for the prostate patient. The nominal DVHs for the constant RBE of 1.1 (solid lines) and for the Wedenberg model (dashed lines) are shown together with their corresponding error bands from the robustness evaluation.

Table 3. Nominal values and ranges of the NTCPs from the robustness evaluation for the three fractionation schedules for one patient of each treatment site.

		NTCP (%)		
ROI and endpoint	RBE model	25 fractions	15 fractions	5 fractions
Lung	RBE 1.1	0.39 (0.37–0.45)	0.39 (0.37–0.45)	0.38 (0.36–0.46)
Grade ≥2 pneumonitis	Wedenberg	0.43 (0.39–0.61)	0.42 (0.38–0.58)	0.40 (0.37–0.54)
	McNamara	0.42 (0.38–0.58)	0.42 (0.38–0.56)	0.40 (0.37–0.53)
		70 fractions	42 fractions	35 fractions
Parotid gland	RBE 1.1	–[a]	11.7 (9.3–14.8)	10.5 (8.3–13.4)
Flow rate <25%	Wedenberg	–[a]	18.0 (11.9–30.8)	15.7 (10.4–26.4)
	McNamara	–[a]	17.0 (11.5–28.2)	15.0 (10.1–24.5)
		39 fractions	20 fractions	5 fractions
Rectum	RBE 1.1	3.3 (1.1–4.3)	3.1 (0.8–5.3)	2.9 (0.3–9.6)
Grade ≥2 rectal toxicity	Wedenberg	5.3 (0.9–20.1)	3.4 (0.5–12.5)	1.2 (0.1–6.7)
	McNamara	6.0 (1.0–21.6)	4.8 (0.6–17.7)	3.2 (0.2–16.6)

[a]Model only applied for fractionation doses between 1.75–2.0 Gy(RBE).

validated model, preferable within clinical trials, before utilized. However, with the method proposed here, one has the opportunity of evaluating the potential impact of the variable RBE and its uncertainty, already at this stage. If the evaluation is not satisfactory, one has the possibility to modify the physical dose and/or the LET_d, in order to search for a solution where e.g., the RBE model of choice agrees better with the constant RBE factor, or a solution where the worst-

case scenarios using the variable RBE still is clinically acceptable.

In conclusion, by disregarding RBE uncertainties, proton plan robustness evaluations may underestimate the variation in the RBE-weighted dose to the CTV and OARs. This study has proposed a method of generating worst-case RBE-weighted dose distributions incorporating the variable RBE and its uncertainties. Such distributions may be of importance to include in comparisons between proton and photon plans, in order to not underestimate the NTCP of critical structures after proton therapy. As the magnitude and direction of the effects are dependent on treatment site, tissue type, biological endpoint, fractionation schedule and RBE model, the analysis should always be done with great care.

Disclosure statement

J.Ö. is part-time employed as an industrial PhD student at RaySearch Laboratories. K.E. is shareholder of RaySearch Laboratories and full-time employed as Chief Science Officer at RaySearch Laboratories. The authors alone are responsible for the content and writing of this article.

Funding

Financial support from the Cancer Research Funds of Radiumhemmet is gratefully acknowledged.

References

[1] Lowe M, Albertini F, Aitkenhead A, et al. Incorporating the effect of fractionation in the evaluation of proton plan robustness to setup errors. Phys Med Biol. 2016;61:413–429.
[2] Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. Phys Med Biol. 2014;59:R419–R472.
[3] Dasu A, Toma-Dasu I. Impact of variable RBE on proton fractionation. Med Phys. 2013;40:011705.
[4] Wedenberg M, Toma-Dasu I. Disregarding RBE variation in treatment plan comparison may lead to bias in favor of proton plans. Med Phys. 2014;41:091706.
[5] Giovannini G, Böhlen T, Cabal G, et al. Variable RBE in proton therapy: comparison of different model predictions and their influence on clinical-like scenarios. Radiat Oncol. 2016;11:68.
[6] Carabe A, España S, Grassberger C, et al. Clinical consequences of relative biological effectiveness variations in proton radiotherapy of the prostate, brain and liver. Phys Med Biol. 2013;58:2103–2117.
[7] Underwood T, Giantsoudi D, Moteabbed M, et al. Can we advance proton therapy for prostate? Considering alternative beam angles and relative biological effectiveness variations when comparing against intensity modulated radiation therapy. Int J Radiat Oncol Biol Phys. 2016;95:454–464.
[8] Ödén J, Eriksson K, Toma-Dasu I. Inclusion of a variable RBE into proton and photon plan comparison for various fractionation schedules in prostate radiation therapy. Med Phys. Forthcoming. [cited 2017 Jan 20]. DOI: 10.1002/mp.12117.
[9] Stokkevåg CH, Fukahori M, Nomiya T, et al. Modelling of organ-specific radiation-induced secondary cancer risks following particle therapy. Radiother Oncol. 2016;120:300–306.
[10] Langendijk JA, Lambin P, De Ruysscher D, et al. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. Radiother Oncol. 2013;107:267–273.
[11] Wedenberg M, Lind BK, Hårdemark B. A model for the relative biological effectiveness of protons: the tissue specific parameter α/β of photons is a predictor for the sensitivity to LET changes. Acta Oncol. 2013;52:580–588.
[12] McNamara AL, Schuemann J, Paganetti H. A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published *in vitro* cell survival data. Phys Med Biol. 2015;60:8399–8416.
[13] Goitein M. Nonstandard deviations. Med Phys. 1983;10:709–711.
[14] Haviland JS, Owen JR, Dewar JA, et al. The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. Lancet Oncol. 2013;14:1086–1094.
[15] Stuschke M, Thames HD. Fractionation sensitivities and dose-control relations of head and neck carcinomas: analysis of the randomized hyperfractionation trials. Radiother Oncol. 1999;51:113–121.
[16] Pedicini P, Caivano R, Fiorentino A, et al. Clinical radiobiology of head and neck cancer: the hypothesis of stem cell activation. Clin Transl Oncol. 2015;17:469–476.
[17] Brenner DJ, Hall EJ. Fractionation and protraction for radiotherapy of prostate carcinoma. Int J Radiat Oncol Biol Phys. 1999;43:1095–1101.
[18] Dasu A, Toma-Dasu I. Prostate alpha/beta revisited - an analysis of clinical results from 14 168 patients. Acta Oncol. 2012;51:963–974.
[19] Fowler JF, Ritter MA, Chappell RJ, et al. What hypofractionated protocols should be tested for prostate cancer? Int J Radiat Oncol Biol Phys. 2003;56:1093–1104.
[20] Joiner MC, van der Kogel A, editors. Basic clinical radiobiology. 4th ed. London: Hodder Arnold; 2009. p. 102–134.
[21] Bentzen SM, Skoczylas JZ, Bernier J. Quantitative clinical radiobiology of early and late lung reactions. Int J Radiat Biol. 2000;76:453–462.
[22] Brenner DJ. Fractionation and late rectal toxicity. Int J Radiat Oncol Biol Phys. 2004;60:1013–1015.
[23] Seppenwoolde Y, Lebesque JV, De Jaeger K, et al. Comparing different NTCP models that predict the incidence of radiation pneumonitis. Normal tissue complication probability. Int J Radiat Oncol Biol Phys. 2003;55:724–735.
[24] Dijkema T, Raaijmakers CPJ, Ten Haken RK, et al. Parotid gland function after radiotherapy: the combined Michigan and Utrecht experience. Int J Radiat Oncol Biol Phys. 2010;78:449–453.
[25] Michalski JM, Gay H, Jackson A, et al. Radiation dose-volume effects in radiation-induced rectal injury. Int J Radiat Oncol Biol Phys. 2010;76:123–129.
[26] Fowler JF. Is there an optimum overall time for head and neck radiotherapy? A review, with new modelling. Clin Oncol. 2007;19:8–22.
[27] Kim M, Stewart RD, Phillips MH. A feasibility study: Selection of a personalized radiotherapy fractionation schedule using spatiotemporal optimization. Med Phys. 2015;42:6671–6678.