Perspectives on the model-based approach to proton therapy trials: a retrospective study of a lung cancer randomized trial

Aimee L McNamara[a,*,†], **David C Hall**[a,†], **Nadya Shusharina**[a], **Amy Liu**[b], **Xiong Wei**[b], **Ali Ajdari**[a], **Radhe Mohan**[b], **Zhongxing Liao**[b], **Harald Paganetti**[a]

[a]Department of Radiation Oncology, Massachusetts General Hospital & Harvard Medical School, Boston, MA

[b]University of Texas MD Anderson Cancer Center, Houston, TX

Abstract

Purpose: The goal of this study was to assess whether a model-based approach applied retrospectively to a completed randomized controlled trial (RCT) would have significantly altered the selection of patients of the original trial, using the same selection criteria and endpoint for testing the potential clinical benefit of protons compared to photons.

Methods and Materials: A model-based approach, based on three widely used normal tissue complication probability (NTCP) models for radiation pneumonitis (RP), was applied retrospectively to a completed non-small cell lung cancer RCT (NCT00915005). It was assumed that patients were selected by the model-based approach if their expected NTCP value was above a threshold of 5%. The endpoint chosen matched that of the original trial, the first occurrence of severe (grade 3) RP.

Results: Our analysis demonstrates that NTCP differences between proton and photon therapy treatments may be too small to support a model-based trial approach for lung cancer using RP as the normal tissue endpoint. The analyzed lung trial showed that less than 19% (32/165) of patients enrolled in the completed trial would have been enrolled in a model-based trial, prescribing photon therapy to all other patients. The number of patients enrolled was also found to be dependent on the type of NTCP model used for evaluating RP, with the three models enrolling 3%, 13% or 19% of patients. This result does show limitations in NTCP models which would affect the success of a model-based trial approach. No conclusion regarding the development of RP in patients randomized by the model-based approach could statistically be made.

Conclusions: Uncertainties in the outcome models to predict NTCP are the inherent drawback of a model-based approach to clinical trials. The impact of these uncertainties on enrollment in

---

[*]Corresponding author at: Department of Radiation Oncology, Massachusetts General Hospital, 55 Fruit St, Boston, 02114, MA, USA. amcnamara2@mgh.harvard.edu.

[†]authors contributed equally to this work.

Conflict of Interest Statement:

The authors report no conflict of interest.

model-based trials depends on the predicted difference between the two treatment arms and on the set threshold for patient stratification. Our analysis demonstrates that NTCP differences between proton and photon therapy treatments may be too small to support a model-based trial approach for specific treatment sites such as lung cancer depending on the chosen normal tissue endpoint.

Introduction

Proton beam therapy (PBT) has the potential to reduce dose to normal tissue, whilst matching the target volume dosimetry achieved by external beam photon therapy [1,2]. Since biological effects are known to increase with dose, PBT is expected to reduce the risk of radiation-related side effects while maintaining rates of local control [3]. In practice, the dose distribution achievable with a given treatment modality depends upon machine parameters (e.g. gantry angles, beam modifiers, etc.), treatment planning strategy, and patient-specific anatomy. Thus, the reduction in toxicity risk offered by PBT and its frequency distribution for a patient population depends not only on the disease site but also on these parameters [4]. Furthermore, the dose-response relationship of organs at risk (OARs) can depend upon clinical risk factors such as age, chemotherapy schedule, and comorbidities [5].

Despite its theoretical underpinnings, there is currently a lack of high-level clinical evidence such as randomized controlled trials (RCTs) for the benefits of PBT [6–9]. A number of RCTs are attempting to rectify this situation, to justify the high cost of PBT (e.g. NCT00875901, NCT02731001) [10]. However, the first completed RCT found no significant difference in radiation pneumonitis (RP) between proton and photon therapy in locally advanced non-small cell lung cancer (NSCLC) [11].

RCTs study the safety and efficacy of a new treatment by comparing outcome to the standard therapy. However in radiotherapy, standardization is not easily defined. RCTs were rarely used for the introduction of new conformal radiation techniques since the ALARA (As Low as Reasonably Achievable) principle was sufficient for justification. However, the high cost of PBT has necessitated the need for clinical evidence. The challenge with RCTs in radiotherapy is that neither the new therapy (PBT) nor the standard therapy (e.g. intensity modulated radiation therapy - IMRT) is identical or consistent between clinics. For example, PBT can use scattering or scanning delivery techniques, similarly photon treatments can vary (IMRT, 3D conformal, etc.) and both are dependent on clinic specific parameters such as planning strategies. Another issue is the ethical prerequisite of RCTs which stipulates that there should always be equipoise. Thus, it would be unethical to include patients that will do better with protons with regard to toxicity in an RCT if the expected benefit is undisputed (such as in pediatric brain tumor patients) [12].

Although RCTs are considered the gold standard of clinical evidence, their appropriateness for investigating the benefits of PBT remains controversial [13–16]. If only a subpopulation is expected to experience a clinically significant reduction in toxicity risk, then perhaps only these patients should be enrolled in the trial. To do otherwise would underestimate the benefits of PBT because the conclusions apply to the entire population. In other words, large-scale RCTs in unselected populations might be unlikely to detect differences but by

enriching the trial with individuals exhibiting a difference in toxicity risk, clinically relevant benefits can be successfully identified for a subpopulation. This is analogous to how biomarkers are used in patient selection when trialing a biologically targeted agent.

To address this concern, a model-based approach to PBT clinical trials was proposed [12,17]. During the patient eligibility evaluation, toxicity risk is estimated using normal tissue complication probability (NTCP) models, as shown in Figure 1. For each patient, the difference between the best proton and photon treatment plan (Δdose) is assessed with respect to the dose-volume parameters in the NTCP-model. To determine the extent the difference in dose translates to a patient-specific reduction in toxicity risk offered by PBT, ΔNTCP, is evaluated. This is defined according to

$$\Delta NTCP = NTCP_{Photons} - NTCP_{Protons}.$$

A positive NTCP indicates that this patient is likely to receive a benefit from PBT compared to X-rays. Patients are enrolled in the trial if their NTCP is above some threshold. The threshold is based on the grade of the endpoint considered. An endpoint grade of 2, 3 and 4–5 should have thresholds greater than 10%, 5% and 2%, respectively [18], based on the assumption that higher toxicity grades have an adverse impact on the patient's quality of life. Multiple toxicities can also be considered by combining NTCP models for different endpoints using an NTCP profile [18]. The threshold restricts the trial cohort to theoretically favorable subpopulations, and thus a model-based RCT should measure a greater effect size. Consequently, a smaller sample size is required to maintain statistical power, reducing the cost of RCTs for PBT.

One current limitation of the model-based approach is that NTCP models derived from proton cohorts for some endpoints currently do not exist. However, a recent study [19], tested the validity of the model-based patient selection approach for PBT by applying photon-derived NTCP models in a head and neck patient population and found that the models remained valid in the PBT patient cohort. Since NTCP models are likely to depend on the radiotherapy technique, this may not be the case for all treatment sites/endpoints. For example, another study recently reported that photon based rectal NTCP models did not predict the observed gastrointestinal morbidity in proton cohorts [20]. Similarly in an oropharyngeal cancer study [21], it was concluded that improved NTCP models were needed for accurate model-based patient selection for PBT.

The aim of our study was to assess the impact of model-based patient selection upon the enrollment of a PBT trial. This was achieved via the retrospective analysis of a completed NSCLC trial. We calculate the number of patients that would be recruited to the proton arm if a model-based approach had been applied. We also investigate retrospectively whether the proton patients selected by the model-based approach had less severe cases of RP compared to those treated with IMRT.

Methods and Materials

Patient cohort:

This study considered patients with locally advanced NSCLC enrolled in a completed RCT (NCT00915005) [11,22]. The trial compared the toxicity and effectiveness of passively scattered proton therapy (PSPT) with that of standard IMRT. Both treatment arms also received concurrent chemotherapy. The primary endpoint was the first occurrence of either severe (grade 3) RP or local failure.

The trial considered enrollment for patients with stage II-IIIB disease, or stage IV disease with a single brain metastasis, or locoregional recurrent tumor after surgical resection that could be treated definitively with concurrent chemoradiation. Additional eligibility criteria included age 18 years but 85 years, Karnofsky Performance Status score 70, and baseline pulmonary function of forced expiratory volume 1 liter. Patients who had received systemic chemotherapy before enrollment, regardless of response, were also eligible.

Radiation treatment planning:

A pair of PSPT and IMRT treatment plans were created for each patient, using four-dimensional computed tomography (4D CT) for motion assessment and target delineation. Both plans were developed using a single set of pre-specified dose-volume constraints, based upon experience with photon therapy (from [11] Suppl. Table S1). The relative biological effectiveness (RBE) was considered constant at 1.1 and the prescribed tumor dose was 74 or 66 Gy(RBE), whichever could be achieved safely within these constraints. Since the development of RP depends on the percentage of the lung volume receiving more than a specific dose (V_{dose}) and the mean lung dose (MLD) [23], patients were eligible for randomization only if both plans satisfied the constraints on the lung V_{20} and MLD. The first 20 patients evaluated were randomly assigned equally to each arm; subsequent patients were assigned with the randomization probability proportion to the 1-year failure rate in each arm.

Initially 272 patients were recruited into the trial, 225 of these were evaluated with comparative IMRT and PSPT plans and 181 produced plans that met the V_{20} and MLD constraints, allowing for a randomized modality assignment. The patients with plans which did not meet the constraints were treated with the acceptable plan (IMRT or PSPT). From the 181 randomly assigned patients, 105 were assigned IMRT and 76 were assigned PSPT. A total of 149 patients were treated with the randomized plan. In our model-based study, a total of 136 randomized and 29 non-randomized patients were included (105 with IMRT and 60 with PSPT)[1].

---

[1]All delivered plans were available however not all undelivered plans were saved on permanent storage after the completion of the original trial. Only patients with both plans from the trial were used in our retrospective study.

Retrospective patient selection:

The model-based approach to PBT trials incorporates an additional model-based patient eligibility criterion, selecting patients expected to receive a clinically significant reduction in toxicity risk from PBT (see Figure 2).

To investigate the impact of this model-based eligibility criterion upon trial enrollment, it was applied retrospectively to the NSCLC patient cohort and the number of patients eligible for randomization was evaluated. The original trial used a combined endpoint, RP and local failure. The model-based approach was originally designed to minimize toxicity effects in healthy tissue instead of evaluating tumor control. For this reason only RP is evaluated in our retrospective study of the model-based approach using NTCP models. The NTCP model parameters used in this study are summarized in Table 1. This retrospective analysis received institutional review board approval.

The NTCP was evaluated using three different NTCP models for RP (see Table 1):

1. Marks et al. (2010) [24]: The QUANTEC logistic regression model of symptomatic RP, which uses mean lung dose as the predictor variable fitted to datasets with grade 1, 2 and 3 with time occurrences ranging from 6 months to no limit.

2. Appelt et al. (2014) [25]: A modified version of the QUANTEC logistic regression model of symptomatic RP. This model is based on the same dataset as the Marks model, but accounts for the odds ratios of various clinical risk factors (see Table 2) and was validated in an independent dataset [25].

3. Tucker et al. (2013) [26]: A Lyman-Kutcher-Burman (LKB) model of severe (grade 3) RP at 12 months, which accounts for volume effects.

A recent study validated NTCP models for NSCLC, including the above models for RP [27]. For our study, the NTCP threshold was chosen to be 5% for RP of grade 3 [12,18] consistent with the original trial. We also investigated a threshold of 10% (for grade 2 RP) for the Marks and Appelt models for comparison, since the occurrence of grade 2 toxicities in the patient dataset we used in this analysis was higher. We did not apply a threshold of 10% for the Tucker model since it is only valid for toxicities of grade 3.

To assess radiation-induced pulmonary toxicity, scored clinical symptoms of RP were extracted from the follow-up medical records based on the National Cancer Institute Common Terminology Criteria for Adverse Events, version 3 (CTCAE v3). RP was scored for 12 months from the completion of radiation therapy. The consensus of up to six clinicians was used for scoring each patient. Clinically symptomatic pneumonitis was defined from grade 2 and higher.

Although the primary endpoint of the original trial was RP, lung cancer patients commonly also experience other toxicities such as esophagitis or cardiac morbidity. Due to the advantageous dosimetry of PBT to spare OARs, such as the heart, we investigate whether other endpoints would be more beneficial for the model-based approach than RP in the trial cohort. NTCP models for esophageal toxicity and cardiac failure were evaluated with

relevant thresholds. Since the original trial only assessed RP as the endpoint, we were only able to include 55 and 60 of the original IMRT and PSPT plans for the heart and esophagus studies, respectively.

Acute esophageal toxicity was assessed with two models (see Table 1):

1. Belderbos et al. (2005) [28]: an LKB NTCP model for grade 2. This model has been externally validated [29].

2. Chapet et al. (2005) [30]: an LKB NTCP model for grade 2–3.

Similarly, two models were used to assess the probability of cardiac failure (see Table 1):

1. Eriksson et al. (2000) [31]: A relative seriality model with parameters based on historical data for long-term cardiac mortality on Hodgkin's disease and breast cancer.

2. Gagliardi et al. (1996) [32]: Model with parameters for cardiac mortality using breast cancer data.

Both heart NTCP models have been applied to lung cancer (e.g. [33]). It has been proposed that toxicity is likely dependent on the portion of the heart that has been irradiated and further studies and data are required to refine the models, especially in patients treated for lung cancer [34].

Table 1 summarizes all models used. The three endpoints and their respective NTCP models are investigated separately for this study and not in combination.

Results

Figure 3 shows the results for the three NTCP models for RP evaluated in the study, for a NTCP threshold of 5% (left panel) and 10% (right panel), corresponding to RP of grade 3 and grade 2, respectively. For each model, the left column shows the NTCP values calculated for both IMRT and PSPT while the right column shows the NTCP. The dotted line in each panel indicates the NTCP threshold. The analysis predicts that out of 165 patients that had both IMRT and PSPT treatment plans; 21, 32, or 5 patients would have been selected in a model-based trial when applying the Marks [24], Appelt [25], and Tucker [26] models for RP with grade 3, respectively. The Marks and Appelt models selected 19 of the same patients. Of the 5 patients selected by the Tucker model, 2 agreed with the Marks model selection and 4 with the Appelt model selection. Only one patient was selected by all three models. When considering a threshold of 10% (RP grade 2), the model-based approach selected 1 and 2 patients for the Marks and Appelt models, respectively. None of the selected patients were common.

As patients were selected for our virtual trial only if they exceeded a certain NTCP threshold, we hypothesized that the patients selected using the model-based approach within the PSPT arm of the original randomized trial, will be patients for whom PBT is associated with better outcome with respect to symptomatic RP. To test this hypothesis, we compared the outcomes for patients treated by IMRT or PSPT according to the trial randomization but retrospectively predicted to benefit from PBT. Table 3 summarizes the results of the

comparison for the three NTCP models to the original trial outcome data. The randomization arm that the patient was originally selected to, as well as the grade of RP the patient experienced post-treatment, is tabulated. We compare the number of proton patients that did not experience RP to those that were selected by the model-based approach to benefit from the treatment. Here we include patients with RP of grade 2 since a recent study found a significant difference in the pulmonary dose response of patients with RP of grade 2 than a grade of 1 [35]. Furthermore, the number of cases of RP 3 in our dataset was too small to draw any significant conclusion. Following our hypothesis, among patients randomized into the IMRT arm, more developed symptoms of RP compared to those assigned to the PSPT arm. This observation did not reach statistical significance because of the limited number of patients enrolled in the virtual trial (p-value of Fisher's exact test > 0.1). The p-values for the Marks, Appelt and Tucker models were 0.65, 1 and 1, respectively.

Figure 4 shows the results for acute esophagitis of grade 2 (left panel) and long-term cardiac mortality (right panel) with a NTCP of 10% and 2%, respectively. For the two NTCP models predicting cardiac toxicity, 44 and 45 out of the 55 patients were selected to benefit from proton treatment. The cardiac models selected 44 patients in common. The esophagitis NTCP models both selected the same 2 patients out of the 60 patients for the model-based trial.

Discussion

Our analysis demonstrates that, based on three widely used NTCP models for RP, dose differences between PSPT and IMRT treatments for a completed lung cancer RCT were too small to support a model-based trial approach on the same patient cohort. These findings may not be applicable to patients treated with a pencil beam scanning technique or for patients with malignancies other than lung cancer. If patients are enrolled in the trial and if their NTCP is above a threshold of 5%, then less than 19% (32/165) of patients enrolled in the completed RCT would have been enrolled in a model-based trial, prescribing IMRT to all other patients.

Uncertainties in outcome models to predict NTCP are an inherent limitation of a model-based approach to clinical trials. The predictive power of NTCP models obtained predominantly from 3D conformal photon treatments for IMRT or PBT treatments is a weakness of this study as it is for the model-based approach in general. The impact of these uncertainties on enrollment in model-based trials depends on the predicted difference between the two treatment arms and on the set threshold for patient stratification. In this study, three different NTCP models for RP of grade 3 each selected a different subset of patients to benefit from protons. The three models used in this study were not mutually congruent with each other, which additionally raises concerns on their reliability. This does suggest that current NTCP models are not strong enough to predict small differences in RP incidence when comparing proton and photon treatments, particularly for use in a model-based approach. It is clear that refinements to NTCP models predicting RP are currently needed. Indeed a recent study using the same dataset [36] found that the effective dose (D_{eff}) to an organ at risk, calculated from a NTCP model, was a better predictor of RP risk than the MLD, since the MLD does not account for high dose effects. Dose response relationships in

the lung could be significantly different between photons and protons due to the underlying energy deposition characteristics not well described by the clinically applied constant RBE of 1.1 [37]. Studies have shown potential elevated RBE values for NSCLC tumors [38] as well as healthy lung [39].

Our primary endpoint for this retrospective study was RP since it was our goal to reassess patient selection of the original trial. NSCLC patients can suffer from multiple other toxicities including heart failure and esophagitis and the associated OARs may show larger differences between proton and photon doses due to lower proton integral doses. We assessed NTCP models for toxicity in the heart and esophagus and found that in the case of heart failure, 65% of the patients would be selected by the model-based approach to benefit from proton therapy, which is higher than in the whole lung toxicity model. However in the case of esophagitis, only 3% of patients would have been selected, which is too low for a statistically significant trial. The treatment plans in this trial were however optimized for MLD and lung V_{20} constraints and not for dose parameters relevant to heart or esophagus endpoints, which could affect these results. The heart NTCP models are based on Hodgkin's and breast cancer data, and although these models have been applied to lung cancer [33], the sub-volume of the heart irradiated may be an important determinate of outcome which needs to be studied for thoracic tumors [34].

Our model-based trial simulation was applied retrospectively to the patient dataset of a completed RCT which may introduce a selection bias in our study. An attempt was made to include all patients from the original trial to this study (225 patients), including those randomized and non-randomized in the trial. This was possible for 165 patients where both the original treatment plans from the RCT were available. Including a subset of the total number of patients may add bias to the study. The original trial protocol used lung dose-volume constraints to determine whether a patient can be randomized or not. In the case where one of the treatment plans did not meet the V_{20} and MLD constraints, then the patient was treated with the plan that did meet these constraints. These constraints assigned 13 patients to PBT (~5.7% of patients). Five of these patients were not included in our model-based trial analysis since their original plans were no longer available. However, this approach does neglect patients who exhibit a large magnitude of NTCP (whether positive or negative).

It is stressed that our study is a hypothesis generating report, which may be used to inform clinical trial design, but should not be interpreted to conclusively determine which patients would benefit from protons vs. photons for locally advanced lung cancer treatment.

In conclusion this study demonstrates, on the example of a recently completed lung cancer RCT, the limitations in the available NTCP models for RP which would affect the reliability of a model-based clinical trial. The model-based approach, based on the reduction of toxicity, is conceptually a correct method to select patients for PBT. However, the Achilles' tendon of the model-based approach is the accuracy and availability of NTCP models for the endpoint(s) of interest in the trial. Model-based trials face an inherent dilemma in that their enrollment is affected by the expected difference between the two treatment arms and the uncertainty in the NTCP model for the chosen endpoint. A model-based approach is thus

effective only if the expected NTCP difference is considerable or if the models have small uncertainties. If one of these factors happens to be true, one might question trial equipoise. On the other hand, such an approach would also lead to large enrollments if the threshold is set very low, thus in turn negating the model-based approach. On the positive side, model-based trial approaches enrich patients for model developments and could therefore contribute to reducing uncertainties in current NTCP models. Stronger models are needed, and future model-based trials will contribute to model improvement.

Acknowledgements

References

[1]. Chang JY, Zhang X, Wang X, Kang Y, Riley B, Bilton S, et al. Significant reduction of normal tissue dose by proton radiotherapy compared with three-dimensional conformal or intensity-modulated radiation therapy in Stage I or Stage III non– small-cell lung cancer. Radiation Oncology Biology 2006;65:1087–96. doi:10.1016/j.ijrobp.2006.01.052.

[2]. Nichols RC, Huh SN, Henderson RH, Mendenhall NP, Flampouri S, Li Z, et al. Proton Radiation Therapy Offers Reduced Normal Lung and Bone Marrow Exposure for Patients Receiving Dose-Escalated Radiation Therapy for Unresectable Stage III Non-Small-Cell Lung Cancer: A Dosimetric Study. Clinical Lung Cancer 2011;12:252–7. doi:10.1016/j.cllc.2011.03.027. [PubMed: 21726825]

[3]. Loeffler JS, Durante M. Charged particle therapy—optimization, challenges and future directions. Nat Rev Clin Oncol 2013;10:411–24. doi:10.1038/nrclinonc.2013.79. [PubMed: 23689752]

[4]. Hall DC, Trofimov AV, Winey BA, Liebsch NJ, Paganetti H. Predicting Patient-specific Dosimetric Benefits of Proton Therapy for Skull-base Tumors Using a Geometric Knowledge-based Method. Radiation Oncology Biology 2017;97:1087–94. doi:10.1016/j.ijrobp.2017.01.236.

[5]. Vogelius IR, Bentzen SM. A literature-based meta-analysis of clinical risk factors for development of radiation induced pneumonitis. Acta Oncol 2012;51:975–83. doi:10.3109/0284186X.2012.718093. [PubMed: 22950387]

[6]. Brada M, Pijls-Johannesma M, De Ruysscher D. Current Clinical Evidence for Proton Therapy. The Cancer Journal 2009;15:319–24. doi:10.1097/PPO.0b013e3181b6127c. [PubMed: 19672149]

[7]. De Ruysscher D, Mark Lodge M, Jones B, Brada M, Munro A, Jefferson T, et al. Charged particles in radiotherapy: a 5-year update of a systematic review. Radiother Oncol 2012;103:5–7. doi:10.1016/j.radonc.2012.01.003. [PubMed: 22326572]

[8]. Allen AM, Pawlicki T, Dong L, Fourkal E, Buyyounouski M, Cengel K, et al. An evidence based review of proton beam therapy: the report of ASTRO's emerging technology committee. Radiother Oncol 2012;103:8–11. doi:10.1016/j.radonc.2012.02.001. [PubMed: 22405807]

[9]. Lievens Y, Pijls-Johannesma M. Health economic controversy and costeffectiveness of proton therapy. Seminars in Radiation Oncology 2013;23:134–41. doi:10.1016/j.semradonc.2012.11.005. [PubMed: 23473691]

[10]. Mishra MV, Aggarwal S, Bentzen SM, Knight N, Mehta MP, Regine WF. Establishing Evidence-Based Indications for Proton Therapy: An Overview of Current Clinical Trials. Int J Radiat Oncol Biol Phys 2017;97:228–35. doi:10.1016/j.ijrobp.2016.10.045. [PubMed: 28068231]

[11]. Liao Z, Lee JJ, Komaki R, Gomez DR, O'Reilly MS, Fossella FV, et al. Bayesian Adaptive Randomization Trial of Passive Scattering Proton Therapy and IntensityModulated Photon Radiotherapy for Locally Advanced Non–Small-Cell Lung Cancer. Jco 2018;36:1813–22. doi:10.1200/JCO.2017.74.0720.

[12]. Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach.

Radiotherapy and Oncology 2013;107:267–73. doi:10.1016/j.radonc.2013.05.007. [PubMed: 23759662]

[13]. Goitein M, Cox JD. Should randomized clinical trials be required for proton radiotherapy? Journal of Clinical Oncology 2008;26:175–6. [PubMed: 18182658]

[14]. Macbeth FR, Williams MV. Proton therapy should be tested in randomized trials. J Clin Oncol 2008;26:2590–1. doi:10.1200/JCO.2008.16.5514.

[15]. Suit H, Kooy H, Trofimov A, Farr J, Munzenrider J, DeLaney T, et al. Should positive phase III clinical trial data be required before proton beam therapy is more widely adopted? No. Radiotherapy and Oncology 2008;86:148–53. doi:10.1016/j.radonc.2007.12.024. [PubMed: 18237800]

[16]. Bentzen SM. Randomized controlled trials in health technology assessment: overkill or overdue? Radiotherapy and Oncology 2008;86:142–7. doi:10.1016/j.radonc.2008.01.012. [PubMed: 18237799]

[17]. Widder J, van der Schaaf A, Lambin P, Marijnen CAM, Pignol J-P, Rasch CR, et al. The Quest for Evidence for Proton Therapy: Model-Based Approach and Precision Medicine. Int J Radiat Oncol Biol Phys 2016;95:30–6. doi:10.1016/j.ijrobp.2015.10.004. [PubMed: 26684410]

[18]. Langendijk JA, Boersma LJ, Rasch CRN, van Vulpen M, Reitsma JB, van der Schaaf A, et al. Clinical Trial Strategies to Compare Protons With Photons. Seminars in Radiation Oncology 2018;28:79–87. doi:10.1016/j.semradonc.2017.11.008. [PubMed: 29735194]

[19]. Blanchard P, Wong AJ, Gunn GB, Garden AS, Mohamed ASR, Rosenthal DI, et al. Toward a model-based patient selection strategy for proton therapy: External validation of photon-derived normal tissue complication probability models in a head and neck proton therapy cohort. Radiotherapy and Oncology 2016;121:381–6. doi:10.1016/j.radonc.2016.08.022. [PubMed: 27641784]

[20]. Pedersen J, Flampouri S, Bryant C, Liang X, Mendenhall N, Li Z, et al. Crossmodality applicability of rectal normal tissue complication probability models from photon-to proton-based radiotherapy. Radiotherapy and Oncology 2019: in press

[21]. Bijman RG, Breedveld S, Arts T, Astreinidou E, de Jong MA, Granton PV, et al. Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. Acta Oncologica 2017;56:1444–50. doi:10.1080/0284186X.2017.1355113. [PubMed: 28828923]

[22]. ClincialTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). Trial of image-guided adaptive conformal photon vs proton therapy, with concurrent chemotherapy, for locally advanced non-small cell lung carcinoma: treatment related pneumonitis and locoregional recurrence; Available from https://clinicaltrials.gov/ct2/show/NCT00915005.

[23]. Wang S, Liao Z, Wei X, Liu HH, Tucker SL, Hu C-S, et al. Analysis of clinical and dosimetric factors associated with treatment-related pneumonitis (TRP) in patients with non–small-cell lung cancer (NSCLC) treated with concurrent chemotherapy and three-dimensional conformal radiotherapy (3D-CRT). Radiation Oncology Biology 2006;66:1399–407. doi:10.1016/j.ijrobp.2006.07.1337.

[24]. Marks LB, Bentzen SM, Deasy JO, Kong F-MS, Bradley JD, Vogelius IS, et al. Radiation dose-volume effects in the lung. Int J Radiat Oncol Biol Phys 2010;76:S70–6. doi:10.1016/j.ijrobp.2009.06.091. [PubMed: 20171521]

[25]. Appelt AL, Vogelius IR, Farr KP, Khalil AA, Bentzen SM. Towards individualized dose constraints: Adjusting the QUANTEC radiation pneumonitis model for clinical risk factors. Acta Oncol 2014;53:605–12. doi:10.3109/0284186X.2013.820341. [PubMed: 23957623]

[26]. Tucker SL, Mohan R, Liengsawangwong R, Martel MK, Liao Z. Predicting pneumonitis risk: a dosimetric alternative to mean lung dose. Int J Radiat Oncol Biol Phys 2013;85:522–7. doi:10.1016/j.ijrobp.2012.03.052. [PubMed: 22580121]

[27]. Thor M, Deasy J, Iyer A, Bendau E, Fontanella A, Apte A, et al. Toward personalized dose-prescription in locally advanced non-small cell lung cancer: Validation of published normal tissue complication probability models. Radiotherapy and Oncology 2019;138:45–51. doi:10.1016/j.radonc.2019.05.011. [PubMed: 31146070]

[28]. Belderbos J, Heemsbergen W, Hoogeman M, Pengel K, Rossi M, Lebesque J. Acute esophageal toxicity in non-small cell lung cancer patients after high dose conformal radiotherapy. Radiotherapy and Oncology 2005;75:157–64. doi:10.1016/j.radonc.2005.03.021. [PubMed: 15890421]

[29]. Dankers FJWM, Wijsman R, Troost EGC, Tissing-Tan CJA, Kwint MH, Belderbos J, et al. External validation of an NTCP model for acute esophageal toxicity in locally advanced NSCLC patients treated with intensity-modulated (chemo)radiotherapy. Radiotherapy and Oncology 2018;129:249–56. doi:10.1016/j.radonc.2018.07.021. [PubMed: 30241789]

[30]. Chapet O, Kong F-M, Lee JS, Hayman JA, Haken Ten RK. Normal tissue complication probability modeling for acute esophagitis in patients treated with conformal radiation therapy for non-small cell lung cancer. Radiotherapy and Oncology 2005;77:176–81. doi:10.1016/j.radonc.2005.10.001. [PubMed: 16256230]

[31]. Eriksson F, Gagliardi G, Liedberg A, Lax I, Lee C, Levitt S, et al. Long-term cardiac mortality following radiation therapy for Hodgkin's disease: analysis with the relative seriality model. Radiotherapy and Oncology 2000;55:153–62. doi:10.1016/S0167-8140(00)00166-3. [PubMed: 10799727]

[32]. Gagliardi G, Lax I, Ottolenghi A, Rutqvist LE. Long-term cardiac mortality after radiotherapy of breast cancer--application of the relative seriality model. Bjr 1996;69:839–46. doi:10.1259/0007-1285-69-825-839. [PubMed: 8983588]

[33]. Woodford K, Panettieri V, Ruben JD, Senthi S. Limiting the risk of cardiac toxicity with esophageal-sparing intensity modulated radiotherapy for locally advanced lung cancers. Journal of Thoracic Disease 2016;8:942–9. doi:10.21037/jtd.2016.03.64. [PubMed: 27162670]

[34]. Gagliardi G, Constine LS, Moiseenko V, Correa C, Pierce LJ, Allen AM, et al. Radiation Dose–Volume Effects in the Heart. Radiation Oncology Biology 2010;76:S77–S85. doi:10.1016/j.ijrobp.2009.04.093.

[35]. Shusharina N, Liao Z, Mohan R, Liu A, Niemierko A, Choi N, et al. Differences in lung injury after IMRT or proton therapy assessed by 18FDG PET imaging. Radiother Oncol 2018;128:147–53. doi:10.1016/j.radonc.2017.12.027. [PubMed: 29352608]

[36]. Tucker SL, Xu T, Paganetti H, Deist T, Verma V, Choi N, et al. Validation of Effective Dose as a Better Predictor of Radiation Pneumonitis Risk Than Mean Lung Dose: Secondary Analysis of a Randomized Trial. Int J Radiat Oncol Biol Phys 2019;103:403–10. doi:10.1016/j.ijrobp.2018.09.029. [PubMed: 30291994]

[37]. McNamara A, Willers H, Paganetti H. Modelling variable proton relative biological effectiveness for treatment planning. Bjr 2019;5:20190334–11. doi:10.1259/bjr.20190334.

[38]. Liu Q, Ghosh P, Magpayo N, Testa M, Tang S, Gheorghiu L, et al. Lung Cancer Cell Line Screen Links Fanconi Anemia/BRCA Pathway Defects to Increased Relative Biological Effectiveness of Proton Radiation. Radiation Oncology Biology 2015;91:1081–9. doi:10.1016/j.ijrobp.2014.12.046.

[39]. Underwood TSA, Grassberger C, Bass R, MacDonald SM, Meyersohn NM, Yeap BY, et al. Asymptomatic Late-phase Radiographic Changes Among Chest-Wall Patients Are Associated With a Proton RBE Exceeding 1.1. International Journal of Radiation Oncology*Biology*Physics 2018;101:809–19. doi:10.1016/j.ijrobp.2018.03.037.

Highlights:

- Aim to assess whether a model-based approach to patient selection for proton therapy trials is always feasible, given modeling uncertainties.

- Model-based trial was applied retrospectively to a completed non-small cell lung cancer (NSCLC) randomized controlled trial.

- Our analysis demonstrates that NTCP differences between proton and photon therapy treatments may be too small to support a model-based trial approach for specific treatment sites such as NSCLC depending on the chosen normal tissue endpoint.
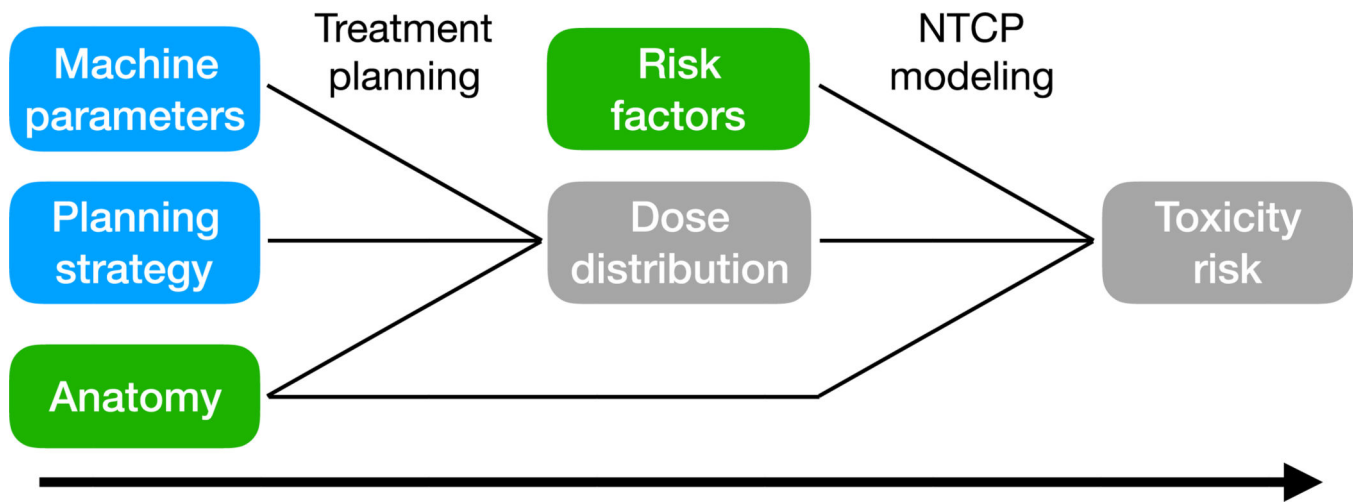
Figure 1.
The process of estimating toxicity risk using a normal tissue complication probability (NTCP) model. Green inputs are associated with the patient. Blue inputs are associated with the treatment arm and should be consistent throughout the trial.
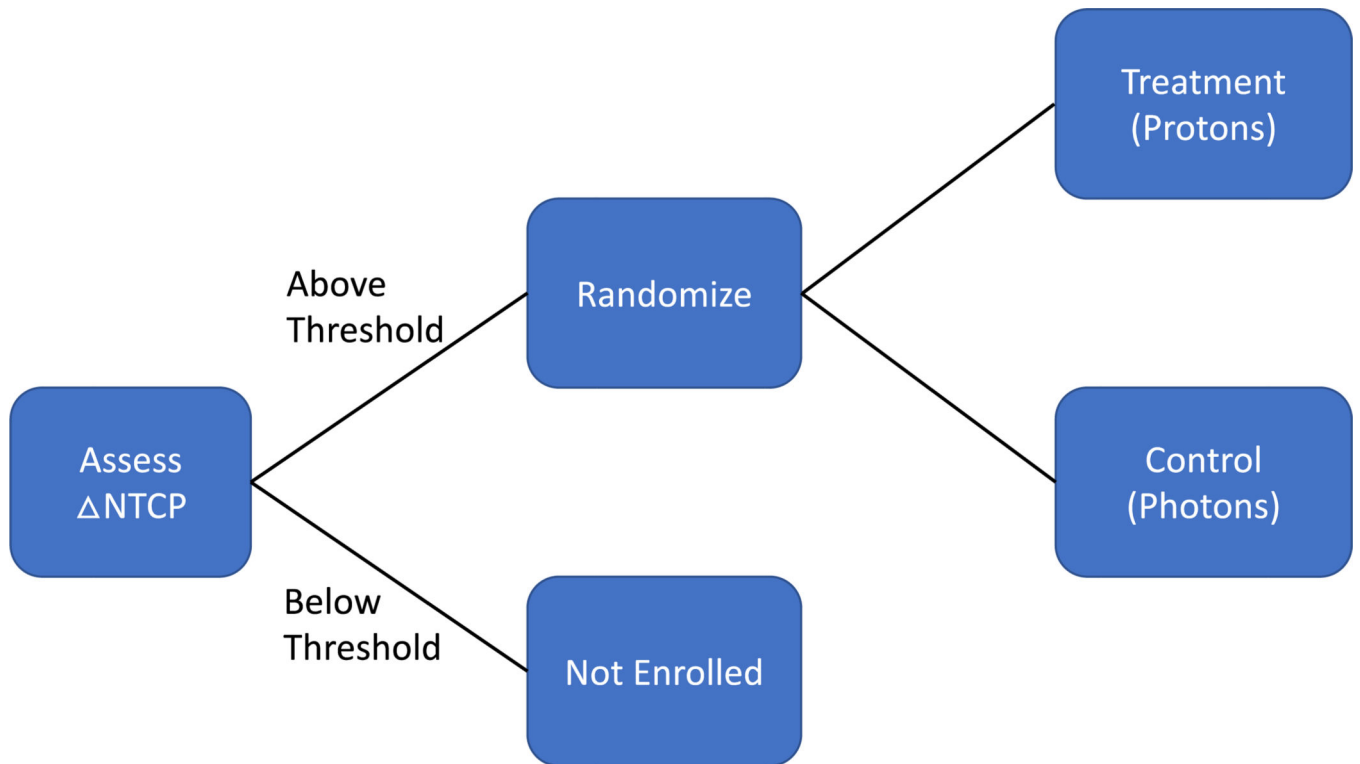
Figure 2.
Enrichment design employed by the model-based approach to proton therapy trials. For our study, the endpoint of interest was radiation pneumonitis (RP) with a grade of 3 and greater and a threshold of 5% was applied.
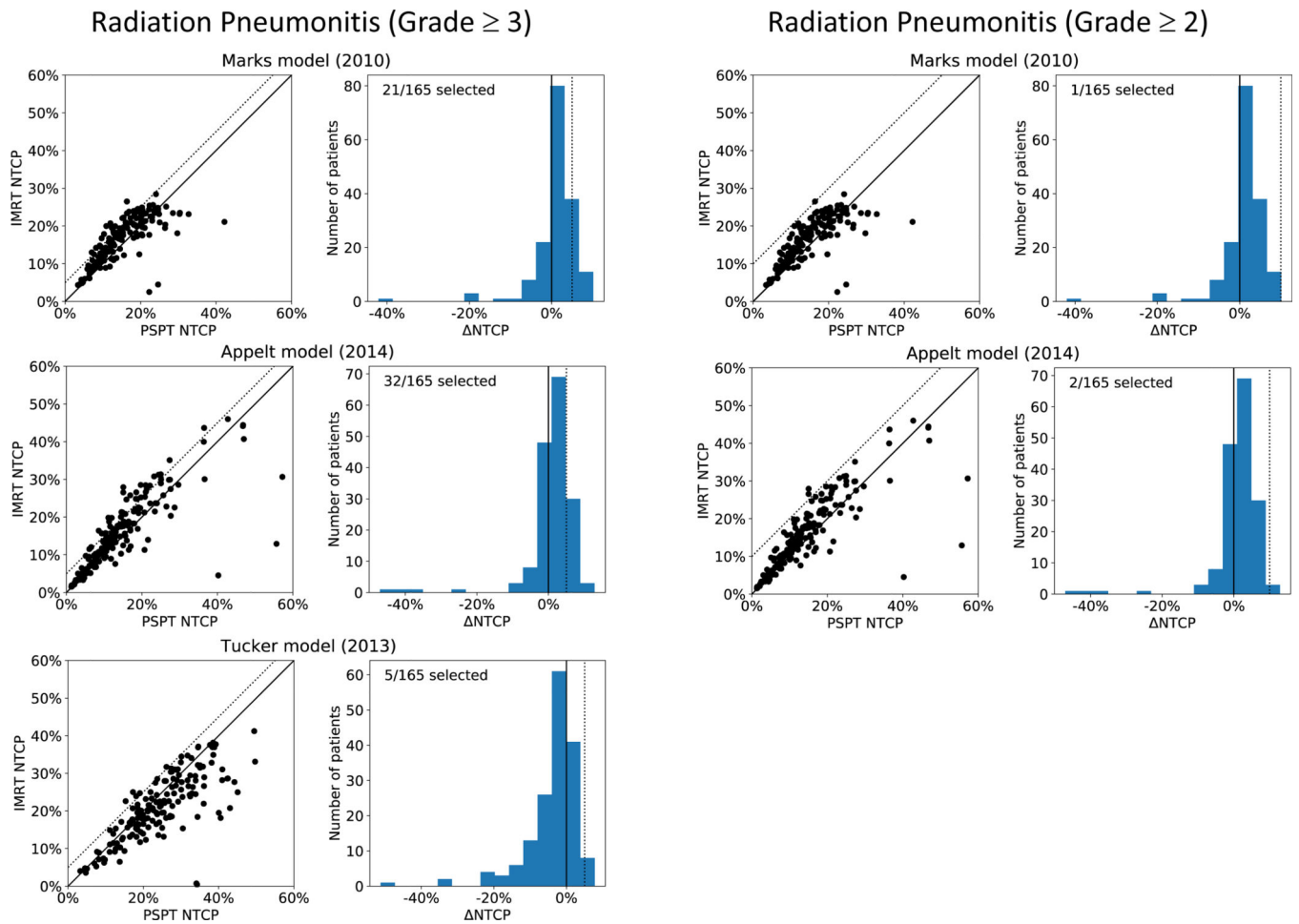
Figure 3.
Results for the three different normal tissue complication probability (NTCP) models used for radiation pneumonitis of grade ≥ 3 (left panel) and grade ≥ 2 (right panel). For each model, the left column shows the calculated NTCP for passive scattering proton therapy (PSPT) and intensity-modulated radiation therapy (IMRT) for all 165 patients (the dotted line indicates the NTCP threshold). The right column shows the respective ΔNTCP. For grade ≥ 3, a threshold of 5% is applied, while for grade ≥ 2 the threshold is set to 10%.
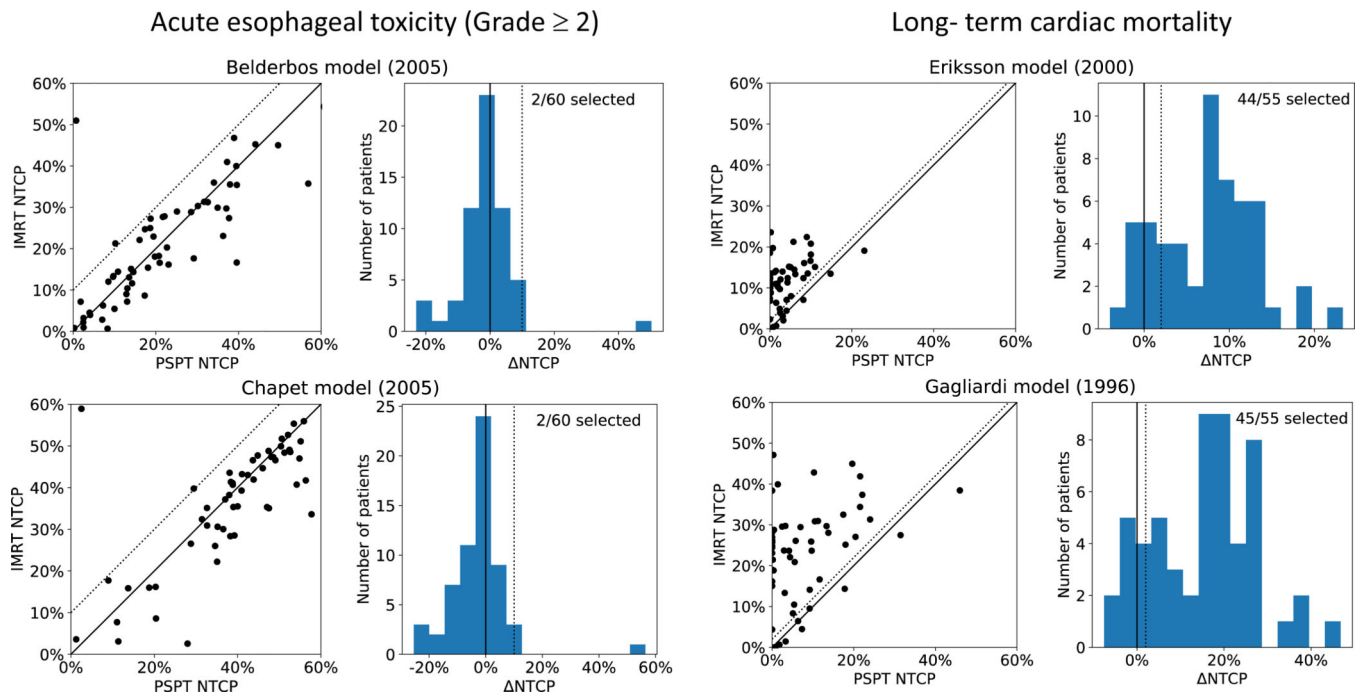
Figure 4.
Results for two other endpoints: acute esophageal toxicity (left panel) and long-term cardiac mortality (right panel). Two normal tissue complication probability (NTCP) models are considered for each endpoint. The left column shows the calculated NTCP for PSPT and IMRT for all patients and the right column shown the respective NTCP. A threshold of 10% was used for esophageal toxicity (grade 2) while for cardiac toxicity a threshold of 2% was used (shown by the dotted lines).

Table 1.

Details of the normal tissue complication probability (NTCP) models used in the analysis. The Lyman-Kutcher-Burman model is abbreviated by LKB.

Model	Endpoint	Severity grade	Type of model	NTCP parameters
Marks et al. (2010) [24]	Radiation pneumonitis	1–3	QUANTEC, logistic fit	$TD_{50} = 30.8$ Gy, $\gamma_{50} = 0.97$
Appelt et al. (2014) [25]	Radiation pneumonitis	1–3	Modified QUANTEC, uses clinical factors	$TD_{50} = 34.4$ Gy, $\gamma_{50} = 1.19$
Tucker et al. (2013) [26]	Radiation pneumonitis	3	Lyman-KutcherBurman	$TD_{50} = 43.2$ Gy, $m = 0.31$, $n = 0.41$
Belderbos et al. (2005) [28]	Acute esophageal toxicity	2	Lyman-Kutcher-Burman	$TD_{50} = 47$ Gy, $m = 0.36$, $n = 0.69$
Chapet et al. (2005) [30]	Acute esophageal toxicity	2–3	Lyman-Kutcher-Burman	$TD_{50} = 51$ Gy, $m = 0.44$, $n = 0.32$
Eriksson et al. (2000) [31]	Long-term cardiac failure		Relative seriality model	$TD_{50} = 70.3$ Gy, $\gamma = 0.96$, $s = 1$
Gagliardi et al. (1996) [32]	Long-term cardiac failure		Relative seriality model	$TD_{50} = 52.3$ Gy, $\gamma = 1.28$ $s = 1$

Table 2.

Odds ratios associated with different clinical risk factors for the development of RP, bracketed values indicate the confidence interval (CI) for 95% [4,25]. Old age was defined with an average cut point of 63 years.

Clinical risk factor	Odds ratio [95% CI]
Pre-existing pulmonary co-morbidity	2.27 [1.25, 4.13]
Mid or inferior tumor location	1.87 [1.26, 2.79]
Current smoker	0.62 [0.43, 0.88]
Former smoker	0.69 [0.47, 1.01]
Old age	1.66 [1.29, 2.13]
Sequential chemotherapy	1.60 [1.11, 2.32]

Table 3.

Treatment outcome with respect to radiation pneumonitis (RP) for patients predicted to benefit from proton therapy treatment using the model-based approach. The number of patients enrolled by the model-based approach are listed. The modality that these patients received in the original trial: intensity-modulated radiation therapy (IMRT) and passively scattered proton therapy (PSPT) is listed and the grade of RP these patients experienced.

Model	Model based approach selected patients	Modality patients received	Number of patients with RP symptoms post-treatment	
			RP of grade 1	RP of grade 2
Marks et al. (2010)	21/165			
		IMRT	7	5
		PSPT	7	2
Tucker et al. (2013)	5/165			
		IMRT	2	2
		PSPT	1	0
Appelt et al. (2014)	32/165			
		IMRT	14	9
		PSPT	5	4