Clinical Trial Strategies to Compare Protons With Photons

Langendijk, Johannes A.; Boersma, Liesbeth J.; Rasch, Coen R. N.; van Vulpen, Marco; Reitsma, Johannes B.; van der Schaaf, Arjen; Schuit, Ewoud

Published in:
Seminars in radiation oncology

DOI:
10.1016/j.semradonc.2017.11.008

Publication date:
2018

ELSEVIER

Seminars in
RADIATION
ONCOLOGY

Clinical Trial Strategies to Compare Protons With Photons

Johannes A. Langendijk, MD, PhD,[*,‖] Liesbeth J. Boersma, MD, PhD,[†,‖] Coen R.N. Rasch, MD, PhD,[†,‖] Marco van Vulpen, MD, PhD,[‡,‖] Johannes B. Reitsma, MD, PhD,[§,‖] Arjen van der Schaaf, PhD,[*,‖] and Ewoud Schuit, PhD[§,‖]

The favorable beam properties of protons can be translated into clinical benefits by target dose escalation to improve local control without enhancing unacceptable radiation toxicity or to spare normal tissues to prevent radiation-induced side effects without jeopardizing local tumor control. For the clinical validation of the added value of protons to improve local control, randomized controlled trials are required. For the clinical validation of the added value of protons to prevent side effects, both model-based validation or randomized controlled trials can be used. Model-based patient selection for proton therapy is crucial, independent of the validation approach. Combining these approaches in rapid learning health care systems is expected to yield the most efficient and scientifically sound way to continuously improve patient selection and the therapeutic window, eventually leading to more cancer survivors with better quality of life.

Introduction

There is a widespread and ongoing discussion on the presumed lack of evidence of protons over photons, which is the most frequently used radiation technique and currently still considered the reference standard for most indications.[1-5] The term "lack of evidence" is often used when results from randomized controlled trials (RCTs) comparing a new treatment modality (eg, protons) with the current standard (eg, photons) is lacking. In this respect, it is important to note that new radiation techniques have rarely been introduced in clinical practice based on the results of RCTs.[6]

Most new radiation techniques are clinically introduced because they allow for better dose conformity (eg, intensity modulated radiotherapy [IMRT], volumetric modulated arc therapy or RapidArc, and protons) and thus better sparing of normal tissues without jeopardizing target dose coverage. To justify the introduction of such techniques in clinical practice, radiation oncologists generally refer to the "ALARA principle," ie, the principle of radioprotection stating that whenever ionizing radiation is applied in humans, animals, or materials, exposure should be "as low as reasonably achievable."[7] As compared to diagnostic imaging, the ALARA principle is considered even more relevant in radiotherapy as the levels of dose exposure administered are markedly higher and more likely to result in clinically apparent acute and late side effects and secondary tumor induction. However, the question arises to what extent the much higher capital and operational costs of proton therapy compared to photon therapy translate into clinically relevant reductions of radiation-induced side effects.

*Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.
†Department of Radiation Oncology (MAASTRO), GROW, School for Oncology and Developmental Biology, Maastricht University Medical Centre+, ZON Proton Therapy Center, Maastricht, The Netherlands.
‡Department of Radiation Oncology, Academic Medical Centre (AMC)—University of Amsterdam, Amsterdam, The Netherlands.
§Holland Proton Therapy Center, Delft, The Netherlands.
‖Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands.
Disclosures: J.A. Langendijk received honoraria from IBA for consulting and advisory role and from IBA Speakers' Bureau, paid to the UMCG Research B.V. The Department of Radiation Oncology has research funding from RaySearch Laboratories (Inst), IBA (Inst), Philips Healthcare (Inst), and Mirada (Inst).
Address reprint requests to Johannes A. Langendijk, MD, PhD, Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713, Groningen, The Netherlands. E-mail: j.a.langendijk@umcg.nl

https://doi.org/10.1016/j.semradonc.2017.11.008

In view of the rising costs of health care, there is a growing societal demand that before the introduction of a new technology in health care such as proton therapy, it must have been shown to be cost-effective, instead of simply referring to the ALARA principle.

On the other side of the spectrum, there are those who propose a direct comparison of protons and photons using the classical approach of an RCT as the one and only acceptable standard of evidence-based medicine, like that used for drug approval. However, there is a growing awareness among different stakeholders that evaluating new technologies with the assessment paradigm used for drug approval may not be the most optimal approach either.[8] RCTs for comparing radiation technologies are much more challenging than for pharmaceutical drugs, owing to the interplay between technological complexity, user skills, local workflows (eg, range and dose verification procedures), additional equipment (eg, treatment planning systems), and learning curve issues, which may all influence the benefits and risks for protons and confuse standardization of the treatment arms. In addition, owing to rapid technological developments in proton therapy, there is a continuous threat that at the time the results of RCTs become available, the outcome will be based on outdated technology and thus not be considered valuable and practice changing. This is not uncommon and has been seen for example for IMRT in head and neck cancer, where results from RCTs that compared IMRT vs 3-dimensional conformal radiotherapy became available when IMRT was already widely used in routine clinical practice.[6] Recently, The Royal Netherlands Academy of Arts and Sciences (KNAW) produced a Foresight Report on the Evaluation of New Technology in Health Care, providing guidelines for research suitable for assessing and inferring the benefits and performance of new technology in health care.[8] They concluded that an RCT is not always the most optimal study design for evaluating the benefit of technology, that a one-size-fits-all approach for the evaluation of medical devices is impossible and that for different types of new applications, different research approaches are required.

In this paper, alternative approaches for an evidence-based sustainable clinical introduction of proton therapy are discussed in addition to methodological problems of RCTs for comparing protons with photons, especially in relation to the eventual introduction of protons into routine clinical practice.

Clinical Applications of Proton Therapy

The favorable beam properties of protons over photons can be translated into clinical benefits in roughly 2 ways.

First, protons can be applied to escalate the dose to the target to improve local tumor control and subsequent overall survival without enhancing additional or unacceptable side effects. According to the Dutch Health Council, dose escalation is the expected future indication for proton therapy in approximately 15% of the cases. This strategy, primarily aiming at improving outcome in terms of efficacy requires the classical approach of an RCT as neither the benefit in terms of improvement of local control and overall survival are known, nor the risks of increasing the dose beyond levels that are normally given to normal tissues in or nearby the target.

Second, protons can be applied to decrease the dose to normal tissues with an equivalent target dose, primarily aiming to prevent acute and late radiation-induced side effects or secondary tumor induction while maintaining similar local tumor control. In the Netherlands, this is the expected application in 85% of the future patients. For this application, clinical validation can be obtained through RCTs under certain specific conditions, but for this strategy alternative methodologies, like the so-called model-based approach, can be considered as well.[1,2]

The Model-Based Approach

The model-based approach is based on the principle that the risk of radiation-induced side effects can be reliably predicted by multivariable normal tissue complication probability (NTCP) models, which are prediction models describing the relationship between dose-volume parameters and the risk on a given side effects.[1] Multivariable NTCP-models consist of at least 1 or more dose-volume parameters either or not in combination with other independent predictors (eg, the addition of concurrent chemotherapy or age).[9-11] The model-based approach can be used to select patients for protons (*model-based selection*); in addition, for the model-based approach it is also essential to continuously and prospectively validate the clinical models for protons (*model-based validation*).

Model-Based Selection

In model-based selection we distinguish 3 steps.

The first step in the model-based approach is to select an NTCP-model or a set of NTCP-models from literature, for acute and late radiation-induced side effects that are considered most relevant (Fig. 1).

In the second step, the dose-volume parameters of the selected NTCP-models are used for optimization of radiotherapy treatment plans, either based on photons or protons (model-based optimization). As prevention of radiation-induced side effects can only be expected when the relevant dose metrics with protons are lower than with photons, the differences between the best proton plan and the best photon plan (Δdose) with respect to the dose-volume parameters in the NTCP-models is assessed by performing a planning comparison study in every single patient.

Step 3 determines to what extent Δdose translates into a difference in complication probability (ΔNTCP) by integrating the results of the planning comparative study into NTCP-models. This final step is necessary as not every Δdose will translate into a clinically relevant ΔNTCP, for example, because the dose with photons already remains under a predefined threshold for a given complication or because of a relatively flat dose-response relationship in the respective Δdose area.

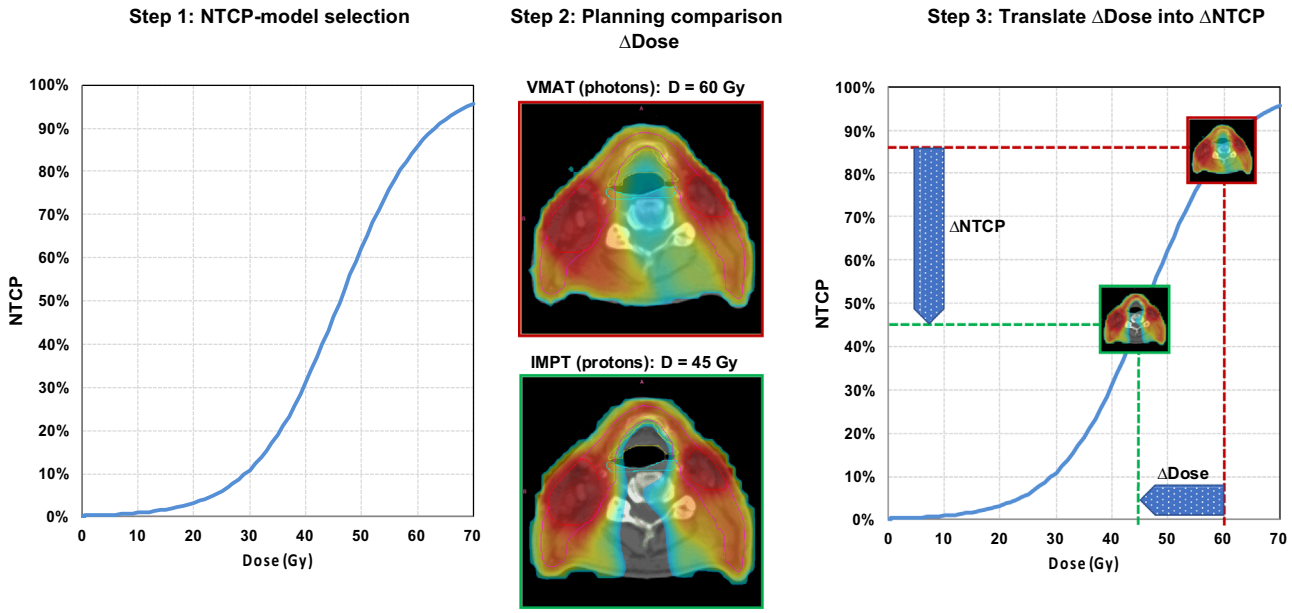

Figure 1 Graphical display of the model-based selection procedure. The first step includes selection of an NTCP-model. Based on the dose-volume parameters included in the selected NTCP-model, the dose distribution is optimized for both techniques (model-based plan optimization) and Δdose is assessed (step 2). Finally, the outcome of step 2 is integrated in the NTCP-model to translate Δdose into ΔNTCP (step 3). (Color version of figure is available online.)

As such, model-based selection is a tool to select patients for proton therapy that are expected to benefit most from this new technology and is a good example of personalized medicine. In addition, a recent study in head and neck cancer patients showed that model-based selection is more likely to be cost-effective than treating all patients with either IMRT or protons.[12] To guarantee uniform decision making, the Dutch Society of Radiation Oncology defined the thresholds for ΔNTCP, which depend on the grading according to the Common Toxicity of Adverse Events version 4.1 (CTCAEv4.0). There is consensus that grade 1 toxicity is not relevant for model-based selection, as this only indicates objective signs of radiation-induced toxicity without any impact for patients or patients' function. In case of grade II, III, and IV-V, the ΔNTCP thresholds that will be used are ≥10%, ≥5%, and ≥2%, respectively, based on the assumption that higher grades of toxicity have more impact on daily function and quality of life.

When multiple NTCP-models for different endpoints are used, the model-based selection procedure eventually results in a so-called ΔNTCP-profile that can considered a biomarker providing quantitative information on the expected benefit of protons compared to protons (Fig. 2). For this situation, the Dutch consensus defined additional thresholds for model-based selection based on ∑ΔNTCP, for example, when 2 NTCP-models are selected, the ∑ΔNTCP threshold is ≥15% with a minimal ΔNTCP threshold of ≥5% for each NTCP-model. Meticulous prospective registration of radiation-induced toxicities including their treatment is required, to determine whether these thresholds indeed lead to a cost-effective indication for proton therapy.

NTCP-Model Selection

The next question is which NTCP-models can be used for model-based selection. The Dutch Platform for Proton Therapy (LPPT) defined the following quality criteria for NTCP-models:

(1) NTCP-models are preferably based on the results of prospective cohort studies, as retrospective assessment of radiation-induced toxicities generally results in underreporting of both the prevalence and the severity of complications. For certain complications, other study designs may be appropriate (eg, nested case control studies), such as complications with very long latency times like cardiac complications after breast cancer radiotherapy or secondary tumor induction.[13,14]
(2) The number of patients and events should be sufficient. In the case of multivariable logistic analysis, a general rule of thumb is that at least 10-15 events are needed for each candidate variable entered in the multivariable model.
(3) NTCP-models preferably are multivariable, considering the effect of dose-volume variables next to other independent predictors to obtain more accurate predictions of the complications risks.
(4) The NTCP-model is presented such that it is possible to calculate the NTCP-values for each individual and for each radiation plan, for example, using an equation, a nomogram, or a graph;
(5) Modern internal validation techniques are applied like bootstrapping or cross-validation.
(6) Information on model performance is available in terms of discrimination and calibration. Discrimination refers to the ability of a model to distinguish patients that will

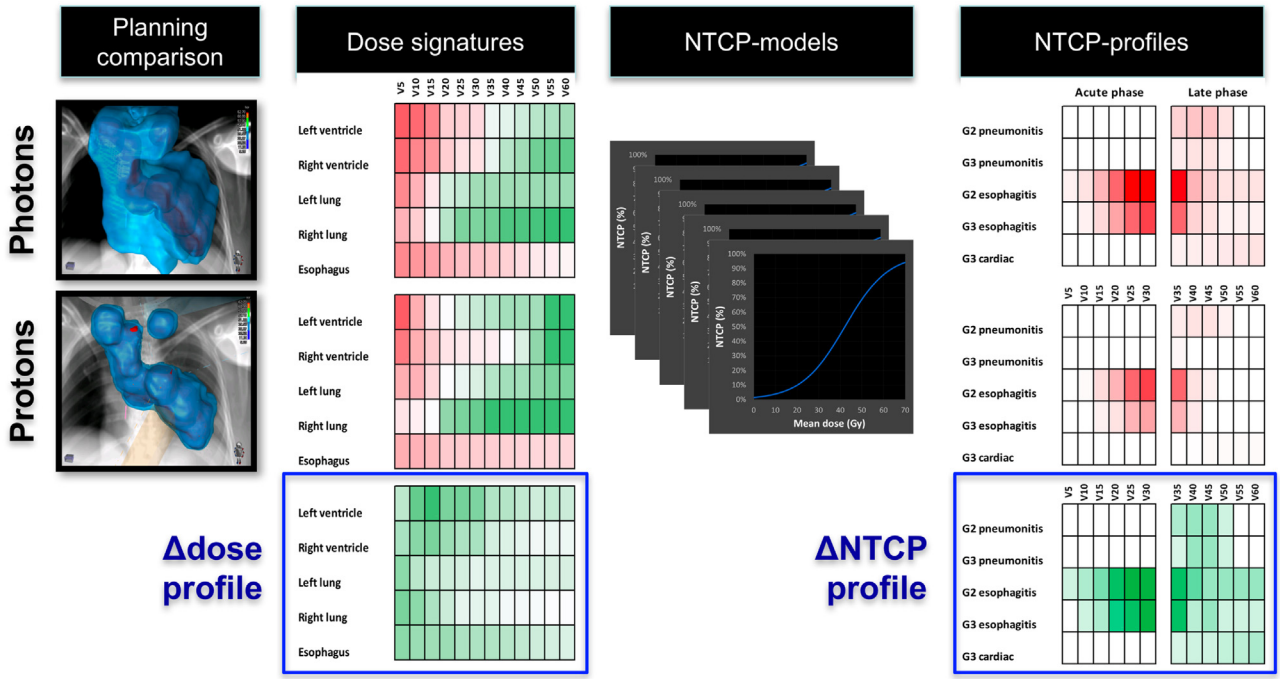

Figure 2 Model-based selection based on ΔNTCP-profile. In case multiple NTCP-models are selected for different toxicity endpoints, the model-based selection procedures results in a Δdose-profile, based on the difference between the dose-volume signatures of photons vs protons which is translated into the ΔNTCP-profile based on a set of multivariable NTCP-models. The ΔNTCP-profile can be considered a biomarker providing quantitative information on a comprehensive set of toxicity risks. The intensity of the green color corresponds linearly with ΔNTCP. ΔNTCP-profiles can be used as a decision support system to decide between photons and protons. The Dutch Society for Radiation Oncology defined standard thresholds for ΔNTCP-profiles based on toxicity grading according to the Common Toxicity Criteria for Adverse Events. ΔNTCP-profiles can also contain patient-reported outcome measures. (Color version of figure is available online.)

or will not develop a given complication, whereas calibration refers to the agreement between predicted and observed outcome.

(7) The NTCP-model is externally validated in an independent data set, preferably in another center. Different levels of external validation can be distinguished and may range from an independent cohort dataset from the same center treated with the same technique (photons) to a cohort from another center with the new technique (protons).

In the Netherlands, the model-based selection procedure for proton therapy has been accepted by the National Health Care Institute (ZiN), which determines whether new types of health care should be included in the basic health insurance package. In 2012 ZiN concluded that the model-based approach can be considered an appropriate evidence-based method to select patients for proton therapy. Consequently, when patients are selected according to the model-based selection procedure, proton therapy is insured care and will be reimbursed by the health care insurers.

Model-Based Validation

The final step in the model-based approach is to clinically validate if a new radiation technique indeed results in less side effects when the relevant dose-volume parameters in the model are decreased by optimizing dose distributions using dose-volume constraints based on the NTCP-models.[1] Recently, Christianen et al[11] were the first to validate the NTCP model for swallowing dysfunction that had been developed in standard IMRT for head and neck cancer, in patients treated with swallowing sparing IMRT.

In a prospective model-based validation study to assess the added value of protons over photons, ΔNTCP is determined in all patients before inclusion, according to the first 3 steps described in the previous paragraph (Fig. 3). Patients are only included if they meet the predefined criteria for ΔNTCP. Eligible patients are treated with the most optimal proton plan, defined as the plan with the highest ΔNTCP compared to the best photon plan.

Subsequently, the average NTCP for both photon plans ($NTCP_{photons}$) and proton plans ($NTCP_{protons}$) of all included patients are calculated, which provides information on the expected toxicity rates with photons and protons and the average ΔNTCP, respectively. Finally, the observed complication rate among those that have been selected for proton therapy is calculated and compared to $NTCP_{photons}$, also referred to as calibration-in-the large.[15] The null-hypothesis, ie, proton therapy does not improve complication rates as compared to photons, is rejected if the observed complication rate with protons is significantly lower than $NTCP_{photons}$. A further test of the observed complication rate against $NTCP_{protons}$ may reveal model deficiencies and possibly trigger

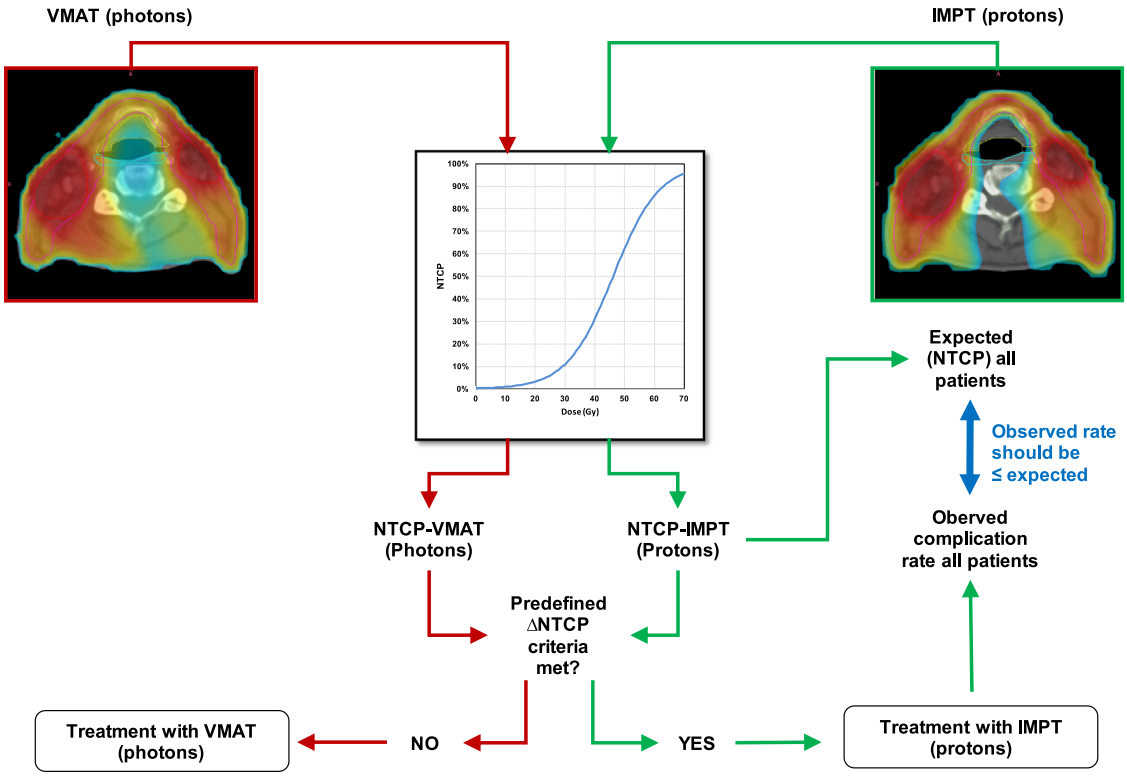

Figure 3 Model-based selection validation of protons. Based on the ΔNTCP-criteria, patients are either selected for photons or protons. After treatment with protons, the observed complication rate is compared to the expected average NTCP for the entire cohort treated with photons. The null hypothesis (ie, no benefit from protons) is rejected if the observed complication rate is significantly lower than the average NTCP for photons. (Color version of figure is available online.)

model updates. Like the design of an RCT, a power analysis should be performed before the study to calculate the number of required patients.

Rapid Learning Health Care

For the model-based approach, it is important to include all patients in a similar data registration program, in which acute and late toxicity preferably in combination with patient-reported outcome measures are prospectively collected at predefined time points, within the framework of a so-called rapid learning health care (RLHC) system[16,17] (Fig. 4). This allows for evaluation of the effect of protons on the primary endpoint (eg, rate of a given late toxicity), but also on a wide spectrum of secondary endpoints (eg, rates of acute and other late toxicities or patient-reported outcome measures) either patient-reported or physician-scored at different time points and can be extended to cohort randomized clinical trials.[16,17] The backbone of the RLHC system is prospective data registration. The data obtained from these registries can be used for the development of multivariable NTCP-models and simultaneous external validation of these models when performed in a multicenter setting. The dose-volume parameters of these models can be used for model-based plan optimization for protons and photons[18,19] and for producing ΔNTCP-profiles to select patients for either protons or photons. All patients are then again subsequently entered in the same prospective data registration program, which allows for model-based validation when patients are treated with protons. These prospectively collected data can also be used to update NTCP-models when needed, for example, by using the closed testing procedure as proposed by Vergouwe et al.[20] The closed testing procedure can be used to decide on the extensiveness of the updating varying from using the original model, through recalibration-in-the-large (ie, re-estimation of the model intercept) and recalibration (ie, re-estimation of the slope and intercept) to the full model revision (ie, re-estimation of all coefficients). It is important to keep the possibility of updating NTCP-models open as some authors found that NTCP-models depend on radiation technology.[21-23] In fact, the RLHC system allows for a continuous improvement of radiotherapy treatment planning and as such will continuously improve outcome of radiotherapy in terms of radiation-induced toxicity.

Randomized Controlled Trials

There may be reasons to prefer an RCT over the model-based approach for the validation of the added value of protons over photons even when the primary aim is to reduce the risk of radiation-induced side effects. RCTs could be considered in case of a complete lack of proper NTCP-models; in case health care authorities demand results of RCTs before reimbursement can even be considered; or, in case there is concern regarding

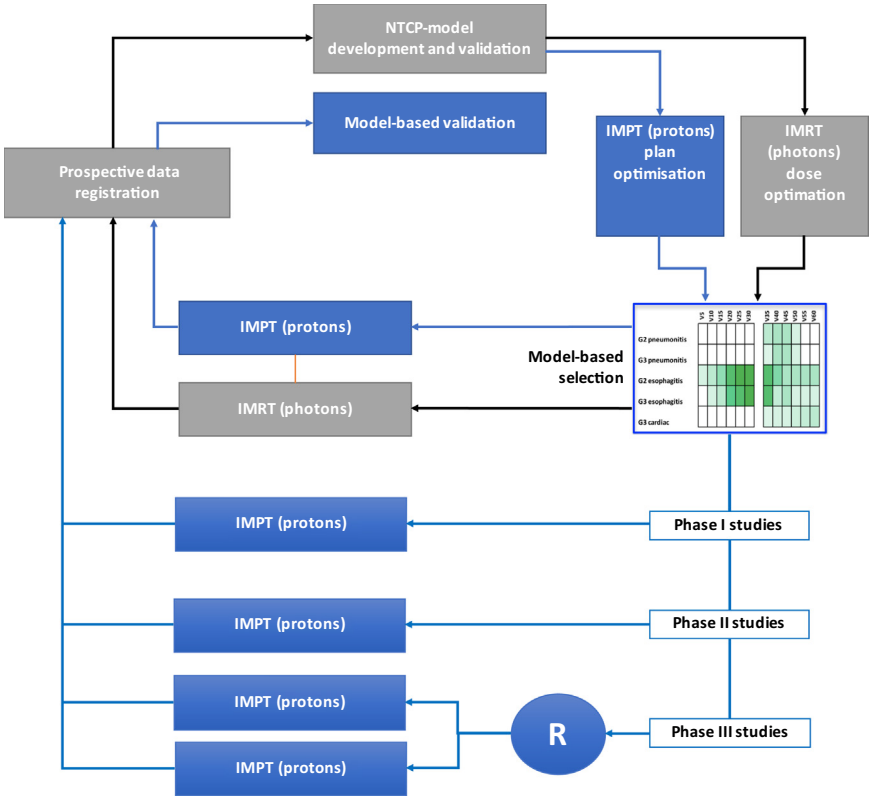

Figure 4 Rapid Learning Health Care system. All patients are subjected to prospective registration program. Data from this registry are used for development and validation of multivariable NTCP-models. The dose-volume metrics in the NTCP-models are used to optimize treatment planning and to produce the ΔNTCP-profile to select patients for either protons or photons. Phase I, II, and III studies can be part of the RLHC system. (Color version of figure is available online.)

decreased tumor control probability (TCP) or increased toxicity other than the primary endpoint when protons are applied, for example, increased dermatological toxicity owing to higher entry dose levels with protons[24] or increased central nervous system toxicity owing to range uncertainties or higher relative biological effectiveness.[25] Like what has been mentioned in the previous paragraphs, in case of conducting RCTs comparing protons with photons, improved outcome with protons in terms of less toxicity can only be expected when the dose to normal tissues is lower (Δdose) and if Δdose is expected to translate into a reduction of NTCP (ΔNTCP).[5] Recently, Widder et al pointed out that even in the case of an RCT, population enrichment (ie, including patients only if a certain level of ΔNTCP can be expected) is required, preferably based on ΔNTCP-profiles if available, or, alternatively, at least based on ΔDose-profiles or other biomarkers. Large and highly powered nonenriched RCTs may yield positive results, even when a large proportion of included cases will not have any benefit based on the absence or clinically irrelevant Δdose- or ΔNTCP-profiles.[5] The consequence of such a "positive" trial could be that when translated into routine clinical practice, too many patients that will not benefit from protons will remain eligible, simply because they meet the nonenriched eligibility criteria of the patients included in the trial. Conversely, owing to dilution of any effect by inclusion of an even larger proportion of patients who predictably will not benefit, a "negative" RCT might thus prohibit treating patients with protons, even the ones with very favorable ΔNTCP-profiles that are likely to benefit from protons.[5]

Another problem with RCTs is that major differences exist among proton therapy centers as illustrated by studies on proton therapy clinical trial credentialing.[26] Similar differences have been found for IMRT.[27,28] Heterogeneity among centers may result from many factors, such as differences in proton therapy delivery equipment, education and training skills, treatment planning systems, treatment planning techniques, center specific workflows, fixation techniques, motion mitigation strategies, and machine and patient-specific quality assurance procedures. In this regard, it is difficult to compare a new heterogeneous technique with IMRT considered as the reference standard, although there is no standard IMRT. This problem becomes increasingly important when results of RCTs are translated into routine clinical practice.

Previous studies showed that the quality of radiation treatment planning and delivery varies widely among institutions and are highly dependent on patient volume.[29,30] Peters et al[30] reported on the results of the quality assurance program of an RCT comparing chemoradiation to chemoradiation with tirapazamine, showing that treatment plans not compliant with the research protocol were associated with significantly worse local control and overall survival rates. Major protocol deficiencies were predominantly associated with lower numbers of patients enrolled in the study, with deficiencies being less common with increasing institutional patient volume.

Similar results were found in the RTOG 0617 lung cancer trial in which patients with locally advanced non-small cell lung cancer were randomly assigned to receive standard vs escalated target dose; treatment at institutions with higher clinical trial accrual volumes was associated with better progression-free and overall survival rates.[29] Patients treated in low volume centers had higher esophageal and heart dose, had more grade 5 events and more termination of radiotherapy owing to severe toxicity.

These results highlight the importance of optimal dose distributions and the need for meticulous quality assurance programs, especially given the potential heterogeneity among proton therapy centers and the possible clinical consequences of these differences on clinical outcome, in terms of efficacy as well as potentially fatal toxicity.

These quality variability issues raise serious questions regarding the generalizability of results obtained from RCTs to routine clinical practice. As the success of treatment and the added value of protons compared to photons depend to a great extent on the performance of individual centers, a positive RCT obtained in a controlled clinical research environment does not a priori improve clinical outcome for individual patients when treated with protons in centers that have not been subjected to similar quality assurance programs. Consequently, even after a successful positive RCT, translation of results from RCTs into routine clinical practice still requires multicenter quality assurance programs, especially for centers that are less experienced with treating eligible patients with the technique investigated. This problem can be addressed in part by integrating quality assurance programs in the RLHC systems.

Cohort Multiple Randomized Controlled Trials

Recently, Relton et al[31] presented another study design: the so-called cohort multiple RCTs (cmRCT). In this design, a specific and large cohort of patients is defined (eg, patients treated with radiotherapy for lung cancer) and followed prospectively, for example, within the framework of an RLHC, with assessment of toxicity and patient-reported outcome measures at fixed time points.[16] This cohort can then be used to test several interventions at the same time. One of the differences with the classical RCT is that patients consent to participate in the prospective data registry and to the possibility to be randomly offered a new intervention (eg, protons instead of photons) when they meet predefined eligibility criteria, whereas the patients in the control group are not approached for further consent. Eventually, the results obtained in the patient cohort treated with the new irradiation technology (eg, protons) can be compared to those obtained in the complete patient cohort treated with the standard radiation technology (eg, photons). The main advantages of cmRCT are: improved accrual rates,[37] the possibility to perform multiple trials within the same cohort, reduced costs for the control arm and the possibility to compare the results to real life practice. This design is particularly useful in case of expensive interventions like protons and high patients' preference to accept the new intervention.[16] Like what has been mentioned for the classical RCT, additional study requirements like population enrichment using model-based selection and advanced quality assurance programs are essential here as well.

Discussion and Conclusion

Different evidence-based clinical trial strategies are available, which can also be applied to investigate whether the use of protons over photons is justified. The choice of trial design depends on several factors, such as the primary study objective (efficacy vs prevention), the availability of high quality multivariable NTCP-models, financial resources and national reimbursement policies. When investigating the added value of protons over photons regarding improvement of local control, RCTs are still the reference standard. When prevention of side effects is the main objective to apply proton therapy, both model-based validation and RCTs can be used, which both have their advantages and disadvantages. Combining both methodologies or with cmRCTs within the framework of RLHC systems is most likely to provide the most complementary evidence-based environment to introduce and validate new radiation modalities like proton therapy.

There is growing awareness in the radiotherapy community that evidence-based selection of patients for new techniques like protons is an unmet need. Some authors reported on other selection methods for proton therapy, aiming at prediction of Δdose, like geometric knowledge based methods in base of skull tumors and head and neck cancer.[31-34] These methods may certainly be helpful in supporting radiation oncologists working in centers where protons are not available to decide, whether to refer patients to a proton therapy center, but may be less valuable in the final decision to treat patients with protons or photons given the earlier mentioned heterogeneity between proton therapy centers. The approach espoused by Lühr et al[35] uses a web-based software tool to support nonproton therapy centers to determine the optimal radiation modality based on patient-specific features.

In addition to selection of patients for proton therapy, model-based selection may become a valuable tool for other radiotherapy applications as well, especially as the availability of proton therapy is expected to remain limited for the next few years. In cases with low tumor control probabilities and unfavorable NTCP-profiles, altered fractionation like hyperfractionation could be considered to increase the therapeutic ratio by increasing local control without enhancing radiation-induced side effects.[36]

Another development is to extend model-based selection by considering not only ΔNTCP but the possible effects on TCP as well, such as in the case of selection of breast cancer patients for proton therapy.[37] It should be noted that the model-based approach as described here assumes that the target dose remains biologically equivalent and that TCP is not affected. To assess the possible effect on TCP, the Department of Radiation Oncology of the University Medical Center Groningen is now implementing a special program for head and neck cancer radiotherapy integrated in the RLHC system: the

QUality AssuraNce TUMor (QUANTUM) program, consisting of 3 major components: (1) A comprehensive quality assurance program for the quality of delineation of target volumes. (2) Automated monitoring of the cumulative target dose-volume parameters during the course of treatment based on repeated CT-scans and dose recalculations and (3) Model-based continuous monitoring of TCP.

At present, one of the problems for the model-based approach is that information on NTCP-models derived from cohorts treated with protons is very limited.[38] As NTCP-models may depend on radiotherapy technique, it is not unlikely that NTCP-models for proton treated patients are different than the NTCP-models based on IMRT treated patients. From a methodological point of view, this is not really a major problem because for both model-based selection and model-based validation, different models for photons and protons can be used simultaneously. The challenge here, however, is to detect possible deviations from the original NTCP-models and the need for model adjustment as soon as possible. In principle, the RLHC system is a powerful tool to detect the need for model adjustment and can be used to continuously evaluate the need for updating NTCP-models using the previously described closed testing procedure.

In conclusion, we have described 3 evidence-based approaches to clinically validate the added value of protons: (1) the model-based approach in which clinical implementation and continuous validation are integrated, (2) the classical RCT approach, and (3) cmRCT. Combining these approaches in RLHC systems is expected to yield the most efficient and scientifically sound way to continuously improve patient selection and the therapeutic window, eventually leading to more cancer survivors with better quality of life.

References

1. Langendijk JA, Lambin P, De Ruysscher D: Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach. Radiother Oncol 107(3):267-273, 2013
2. Lievens Y, Pijls-Johannesma M: Health economic controversy and cost-effectiveness of proton therapy. Semin Radiat Oncol 23(2):134-141, 2013
3. De Ruysscher D, Belderbos J, Reymen B: State of the art radiation therapy for lung cancer 2012: A glimpse of the future. Clin Lung Cancer 14 (2):89-95, 2013
4. Bekelman JE, Hahn SM: Reference pricing with evidence development: A way forward for proton therapy. J Clin Oncol 32(15):1540-1542, 2014
5. Widder J, van der Schaaf A, Lambin P: The quest for evidence for proton therapy: Model-based approach and precision medicine. Int J Radiat Oncol Biol Phys 95(1):30-36, 2016
6. Nutting CM, Morden JP, Harrington KJ: PARSPORT trial management group. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): A phase 3 multicentre randomised controlled trial. Lancet Oncol 12(2):127-136, 2011
7. Uffmann M, Schaefer-Prokop C: Digital radiography: The balance between image quality and required radiation dose. Eur J Radiol 72 (2):202-208, 2009
8. Moons K, Meier I, Bossuyt P, et al: Evaluation of New Technology in Health Care. In Need of Guidance for Relevant Evidence. Amsterdam: KNAW, 2014. pdf available on ⟨www.knaw.nl⟩
9. Wopken K, Bijl HP, van der Schaaf A, et al: Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol 113(1):95-101, 2014
10. Beetz I, Schilstra C, van der Schaaf A: NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: The role of dosimetric and clinical factors. Radiother Oncol 105(1):101-106, 2012
11. Christianen ME, van der Schaaf A, van der Laan HP: Swallowing sparing intensity modulated radiotherapy (SW-IMRT) in head and neck cancer: Clinical validation according to the model-based approach. Radiother Oncol 118(2):298-303, 2016
12. Ramaekers BL, Grutters JP, Pijls-Johannesma M: Int J Radiat Oncol Biol Phys 85(5):1282-1288, 2013
13. Darby SC, Ewertz M, McGale P: Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med 368(11):987-998, 2013
14. Grantzau T, Overgaard J: Risk of second non-breast cancer among patients treated with and without postoperative radiotherapy for primary breast cancer: A systematic review and meta-analysis of population-based studies including 522,739 patients. Radiother Oncol 121(3):402-413, 2016
15. Steyerberg EW, Vickers AJ, Cook NR: Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology 21(1):128-138, 2010
16. Lambin P, Zindler J, Vanneste B: Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. Acta Oncol 54(9): 1289-1300, 2015
17. Lambin P, Roelofs E, Reymen B: 'Rapid Learning health care in oncology'—An approach towards decision support systems enabling customised radiotherapy. Radiother Oncol 109(1):159-164, 2013
18. Kierkels RG, Wopken K, Visser R: Multivariable normal tissue complication probability model-basedtreatment plan optimization for grade 2-4 dysphagia and tube feeding dependencein head and neck radiotherapy. Radiother Oncol 121(3):374-380, 2016
19. Kierkels RG, Korevaar EW, Steenbakkers: Direct use of multivariable normal tissue complication probability models in treatment plan optimisation for individualised head and neck cancer radiotherapy produces clinically acceptable treatment plans. Radiother Oncol 112(3):430-436, 2014
20. Vergouwe Y, Nieboer D, Oostenbrink R: A closed testing procedure to select an appropriate method for updating prediction models. Stat Med 36 (28):4529-4539, 2017
21. Beetz I, Schilstra C, van Luijk P: External validation of three dimensional conformal radiotherapy based NTCP models for patient-rated xerostomia and sticky saliva among patients treated with intensity modulated radiotherapy. Radiother Oncol 105(1):94-100, 2012
22. Troeller A, Yan D, Marina O, Schulze D: Comparison and limitations of DVH-based NTCP models derived from 3D-CRT and IMRT data for prediction of gastrointestinal toxicities in prostate cancer patients by using propensity score matched pair analysis. Int J Radiat Oncol Biol Phys 91 (2):435-443, 2015
23. Christianen ME, Schilstra C, Beetz I: Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. Radiother Oncol 105(1):107-114, 2012
24. Galland-Girodet S, Pashtan I, MacDonald SM: Long-term cosmetic outcomes and toxicities of proton beam therapy compared with photon-based 3-dimensional conformal accelerated partial-breast irradiation: A phase 1 trial. Int J Radiat Oncol Biol Phys 90(3):493-500, 2014
25. Giantsoudi D, Sethi RV, Yeap BY: Incidence of CNS injury for a cohort of 111 patients treated with proton therapy for medulloblastoma: LET and RBE associations for areas of injury. Int J Radiat Oncol Biol Phys 95(1): 287-296, 2016
26. Taylor PA, Kry SF, Alvarez P: Results from the imaging and radiation oncology core Houston's anthropomorphic phantoms used for proton therapy clinical trial credentialing. Int J Radiat Oncol Biol Phys 95(1): 242-248, 2016
27. Das IJ, Andersen A, Chen ZJ: State of dose prescription and compliance to international standard (ICRU-83) in intensity modulated radiation therapy among academic institutions. Pract Radiat Oncol 7(2): e145-e155, 2017

28. Fairchild A, Langendijk JA, Nuyts S, et al: Quality assurance for the EORTC 22071-26071 study: Dummy run prospective analysis. Radiat Oncol 9:248, 2014
29. Eaton BR, Pugh SL, Bradley JD: Institutional enrolment and survival among NSCLC patients receiving chemoradiation: NRG Oncology Radiation Therapy Oncology Group (RTOG) 0617. J Natl Cancer Inst 108(9):1-8, 2016
30. Peters LJ, O'Sullivan B, Giralt J: Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. J Clin Oncol 28(18):2996-3001, 2010
31. Relton C, Torgerson D, O'Cathain A, et al: Rethinking pragmatic randomised controlled trials: Introducing the 'cohort multiple randomized controlled trial' design. BMJ 340:c1066, 2010
32. Vickers AJ: Clinical trials in crisis: Four simple methodologic fixes. Clin Trials 11(6):615-621, 2014
33. Hall DC, Trofimov AV, Winey BA: Predicting patient-specific dosimetric benefits of proton therapy for skull-base tumors using a geometric knowledge-based method. Int J Radiat Oncol Biol Phys 97(5): 1087-1094, 2017
34. Delaney AR, Dahele M, Tol JP, Kuijper IT, Slotman BJ, Verbakel WFAR: Using a knowledge-based planning solution to select patients for proton therapy. Radioth Oncol 124(2):263-270, 2017
35. Lühr A, Löck S, Roth K, Helmbrecht S: Concept for individualized patient allocation: ReCompare—remote comparison of particle and photon treatment plans. Radiat Oncol 9:59, 2014
36. Azria D, Lapierre A, Gourgou S: Data-based radiation oncology: Design of clinical trials in the toxicity biomarkers era. Front Oncol 7:83, 2017
37. Stick LB, Yu J, Maraldo MV, et al: Joint estimation of cardiac toxicity and recurrence risks after comprehensive nodal photon versus proton therapy for breast cancer. Int J Radiat Oncol Biol Phys 97(4):754-761, 2017
38. Blanchard P, Wong AJ, Gunn GB: 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. Radiother Oncol 121(3):381-386, 2016