Time is NTCP: should we maximize patient throughput or perform online adaptation on proton therapy systems?

Authors:

E. Borderías-Villarroel[1], A. Barragán Montero[1], E. Sterpin[1,2,3]

[1]*UCLouvain, Institut de recherche expérimentale et clinique, Molecular Imaging and Radiation Oncology (MIRO) Laboratory, Brussels, Belgium*

[2]*KULeuven, Department of Oncology, Laboratory of external radiotherapy, Leuven, Belgium*

[3]*Particle Therapy Interuniversity Center Leuven – PARTICLE, Leuven, Belgium*

Corresponding author

Edmond Sterpin: edmond.sterpin@uclouvain.be

Abstract

Background: Compared to conventional radiotherapy (XT), proton therapy (PT) may improve normal tissue complication probabilities (NTCP). However, PT typically requires higher adaptation rates due to an increased sensitivity to anatomical changes. Systematic online adaptation may address this issue, but it requires additional replanning time, decreasing patient throughput. Therefore, less patients would benefit in such case from PT for a given machine capacity, with results in worse NTCP

Aim: To investigate the trade-off between PT patient throughput and NTCP gain as a function of the time needed for adaptation.

Methods: A retrospective database of 14 lung patients with two repeated 4DCTs was used to compare NTCP values between XT and PT for $NTCP_{2ym}$ (2-year mortality), $NTCP_{dysphagia}$ and $NTCP_{pneumonitis}$. Four scenarios were considered for PT: no adaptation using clinical robustness parameters (4D robust optimization, 3% range error and PTV-equivalent setup errors); systematic online adaptation with clinical robustness parameters; setup errors reduced to 4 mm and to 2 mm. Dose was accumulated on the planning CT. The number of patients treated with PT depended on the extra time needed for adaptation, assuming an 8-hours capacity (assuming 14 patients a day; thus minimum 34.2 minutes per treatment session if there is no or instantaneous adaptation).

Results:
Baseline NTCP gains (PT against XT without adaptation) equaled 6.9%, 6.1%, and 7.7% for $NTCP_{2ym}$, $NTCP_{dysphagia}$ and $NTCP_{pneumonitis}$, respectively. Using instantaneous online adaptation and setup errors of 2 mm, the overall gains were then 10.7%, 13.6% and 12.4%. Taking into account loss of capacity, 13.7 min was the maximum extra-time allowed to complete adaptation and maintain an advantage on all three metrics for the 2-mm setup error scenario.

Conclusion: This study highlights the critical importance of keeping short online adaptation times when using systems with limited capacity like PT.

1. Introduction

Daily online adaptive radiation therapy (OART) has become a reality in recent years with the arrival of MR-Linacs [1] and the Ethos Varian System [2]. OART offers the possibility of accounting for changes in patient anatomy and tumor geometry that otherwise can compromise the desired target coverage and organs-at-risk (OARs) sparing. OART has demonstrated a greater ability to spare OARs and ensure target coverage for head and neck [3], gastrointestinal [4], pelvic [2,5,6], and finally ultracentral [7] and locally advanced lung cancers [8]

The potential of daily online adaptation could be even greater when applied in proton therapy due to its higher adaptation rate [9]. Safety margins and (anatomical) robust optimization, are typically used to mitigate the impact of treatment uncertainties and reduce the need for adaptation due to the inter-fractional anatomic variation [10]. However, both strategies come at the cost of normal tissue irradiation and even though, for some cases, adaptation is still necessary to ensure target coverage [11,12]. As reported in a recent survey of practice patterns for adaptive particle therapy, from 68 clinically operational centers consulted worldwide, 84% had implemented offline adaptation strategies, while none of them relied on daily online re-planning [9]. Nevertheless, daily online re-planning was determined to be the method of choice at the 10-year horizon, requiring automated and quick workflows to compensate for inter-fractional variations in a realistic online setting.

With current technology, all strategies for online adaptive proton therapy (OAPT) proposed in the literature require additional time compared to non-adaptive treatments. For OAPT strategies not relying on contour re-delineation (i.e. dose restoration [13,14]), this additional time comes from re-optimization and quality assurance (QA) of the adapted treatment plan. For alternative OAPT strategies based on automatic contouring, the time for contour verification (and/or correction) by the physician should be also summed up. For OART in the pelvic region, McComas et al. showed that around 16 additional minutes were required for each adaptive fraction, with an average minimum cost of $103.58 [15]. This study also pointed out that approximately 3 patients could be treated with image-guided radiation therapy (IGRT) for every 1 OART treatment. It also concludes that each adapted fraction requires more staffing while it also delays the completion of other tasks and increases resource utilization. Therefore, before the progressive implementation of OAPT technologies, it urges quantifying the clinical gain of adaptation while considering the cost of time and human resource requirements.

The increased time per fraction experienced by OAPT could lead to a reduction in the treatment capacity of proton therapy facilities, which are limited due to the upfront and running costs associated to that technology. Therefore it becomes imperative to justify the clinical gains achieved through OAPT relative to the resulting capacity loss as it could result in patients delaying their treatment or receiving photon radiation therapy instead.

To estimate real costs (e.g. in, euros or dollars), an assessment must be made of (1) the actual cost of each modality and (2) the societal gains linked to treatment outcome improvements, possibly quantified via Quality-Adjusted Life Years (QALY). Both aspects, fixed and variable costs, pose challenges for accurate quantification [16,17]. Alternatively, health economics considerations based on normal tissue complication probabilities (NTCP) can be considered to study cost-effectiveness. For instance, thresholds for ΔNTCP between two modalities (protons and photons) are used for patient referral [18,19], as it is the case for the Dutch national protocol.

Following this idea, we present a novel approach to compare the merits of treatment adaptation on a proton therapy machine with maximizing its capacity. Unlike traditional cost-based estimations, our method relies on quantifying variations on NTCP to bypass the estimation of actual costs. The objective of this study is therefore to quantify the impact of increased time per fraction on the NTCP for a dataset of 14 lung cancer patients undergoing OAPT. By quantifying the ΔNTCP as a function of the time needed to perform an adaptation, we aim to provide insights on how to optimize lung cancer patient outcomes treated with proton therapy, effectively balancing the benefits of OAPT against the constraints of treatment capacity.

2. Material and methods

2.1. Patient database

This study included a cohort of fourteen lung cancer patients previously treated with tomotherapy (XT) that were retrospectively re-planned with intensity-modulated proton therapy (IMPT) strategies. The dataset comprised a 4D-planning-CT scan and two subsequent repeated 4D-CT scans (rCT1 and rCT2), acquired at regular intervals of 2 and 4 weeks during the course of radiotherapy. Exceptionally, for patient 9, rCT2 was not available, so a total of 27 rCTs were available. All rCTs contained manually delineated contours. The patient cohort represented various tumor sizes (79-433 cm3), motion amplitudes (1.1-11.1 mm), tumor locations, and an average tumor shrinkage of the CTV volume of 16.7% for rCT1 and 19.1% for rCT2, with only patients 1 and 7 experiencing tumor growth in rCT1 (see supplementary material SM1). The retrospective data

collection was approved by the ethical committee of Cliniques Universitaires St. Luc and UCLouvain.

2.2. Initial treatment planning on planning-CTs

The XT plans were optimized with the TomoTherapy treatment planning system using a safety planning target volume (PTV) margin derived from Van Herk's margin recipe ($M_{PTV} = 2.5\Sigma_{total} + 0.7\sigma_{total}$) accounting for total systematic (Σ) and random (σ) uncertainties [20,21]. On the other hand, IMPT plans were generated using worst-case 4D robust optimization [22] because of their sensitivity to the uncertainties of the position of the Bragg peak and the failure of the static-dose cloud approximation [23]. The magnitude of range and setup robustness parameters can be approximated as a function of systematic and random uncertainties to ensure adequate CTV coverage [24,25]. Therefore, to generate an IMPT equivalent to the XT plans, the value of the PTV margin in each direction was translated into a non-isotropic setup error for robust optimization. For cases including treated lymph nodes, the calculated setup errors were 7.71 mm (left-right), 8.25 mm (anterior-posterior) and 9.73 mm (inferior-superior) (see supplementary material SM2). To speed up the optimization of the IMPT plan, with setup errors higher than 5 mm, the optimization trick described by Badiu et al. was implemented [26]. The full non-isotropic setup error was divided into two components: 5 mm were included on the RayStation robust optimization while the rest of the setup error was accounted for with a CTV expansion. For instance, for a setup error of 7.71 mm in the sagittal direction, a CTV expansion of 2.71 mm was robustly optimized with a setup error of 5 mm (7.71 mm = 2.71 mm in the margin + 5 mm setup error in the robust optimization).

The total number of optimization scenarios was 84: 7 (± variable setup error (mm) in x,y,z directions, and the nominal scenario) × 3 (range error: ±3%, 0%) × 4 (breathing phases mid-position-CT (MidP-CT) used as planning CT [27], together with end-inhale, end-exhale and mid-ventilation breathing phases).

The higher the robustness parameters, the less adaptation required. However, the implementation of an OAPT workflow opens the possibility to reduce margins or, as an equivalent, decreasing robustness parameters (i.e. setup error) compared to non-adaptive workflows. Therefore, we investigate the NTCP-time correlation for two additional plans with reduced setup errors of 4 and 2 mm. The 4 mm setup error, reflecting an intermediate approach, considers both inter- and intra-fraction uncertainties, particularly important for mobile targets, while acknowledging the ongoing research to validate margin or setup error reduction without compromising tumor control [28,29]. Even in optimistic scenarios, with daily imaging and high-speed optimization, a 2 mm setup error should be considered to offset residual uncertainties

related to isocenter position, delineation errors and limited intra-fraction uncertainties [30].

After the planning stage on pCT, 4 radiation therapy plans (one XT and three IMPT) were available:

1. **XT_{pCT}**: The clinical tomotherapy plan, using the previously detailed CTV-to-PTV non-isotropic margins between 7 and 9 mm.
2. **IMPT-$Clinic_{pCT}$**: A clinical IMPT plan robustly optimized using a non-isotropic setup error to generate a "PTV-like" equivalence to the XT plan (1).
3. **IMPT-$4mm_{pCT}$**: An IMPT plan accounting for a reduced setup error of 4 mm
4. **IMPT-$2mm_{pCT}$**: An IMPT plan with a residual setup error of 2mm

2.3. Adaptation on repeated-CTs, dose mapping and accumulation

The objective is to assess whether the ΔNTCP in several OAPT settings (considering increased time per fraction and therefore the potential loss proton treatment capacity) is bigger than the ΔNTCP gained from passing from a XT to an IMPT treatment plan without considering adaptation. This will be accomplished without setup error modifications to the clinical plan (OAPT-Clinic) but also when applying two levels (2 and 4 mm) of setup error reduction. To do so, a total of 5 different workflows had to be simulated: (1) the XT workflow (represented by the reference XT dose), (2) the not-adapted workflow and the (3) OAPT-Clinic, (4) OAPT-4mm and (5) OAPT-2mm workflows (see Figure 1).

For the XT plans, we assumed that the planned XT doses on the pCT were representative of the actual dose delivered to the patient. To simulate the not adapted plans straight dose calculation of the IMPT-Clinic plans was performed on all rCTs (Not-$Adapted_{rCT1}$ and Not-$Adapted_{rCT2}$, red boxes in Figure 1).

For the three OAPT workflows (OAPT-Clinic, OAPT-4mm and OAPT-2mm), adapted versions of the IMPT plans ($Adapted_{rCT1(2)}$ IMPT-Clinic, $Adapted_{rCT1(2)}$ IMPT-4mm and $Adapted_{rCT1(2)}$ IMPT-2mm) were generated. As our objective was to simulate a high-quality OAPT workflow, the re-planning was done offline to avoid uncertainties introduced by the OAPT workflow itself. The same objectives used for initial treatment planning on pCT were used for re-planning on rCTs and fine-tuning was applied when necessary to ensure target coverage and avoid hot spots.

Eventually, a total of 12 dose distributions were available for each patient: 4 doses on the pCT (XT_{pCT}, IMPT-$Clinic_{pCT}$, IMPT-$4mm_{pCT}$ and IMPT-$2mm_{pCT}$), 2 not adapted doses (rCT1,rCT2) and 6 adapted doses, corresponding to OAPT-Clinic, OAPT-4mm and OAPT-

2mm in rCT1 and rCT2. These doses were used to simulate the 5 workflows described in Figure 1.

To approximate the total dose delivered over 30 fractions for the not adapted and the OAPT workflows, dose mapping and accumulation (DMA) was applied. The dose mapping from rCTs to the pCT was done using the deformable image registration (DIR) algorithm of RayStation 12B (RaySearch Laboratories, Sweden). As shown in Figure 1, the total accumulated doses for four different plans - the not adapted IMPT-Clinic (NA-Clinic), adapted IMPT-Clinic (OAPT-Clinic), the adapted IMPT-4mm (OAPT-4mm) and the adapted IMPT-2mm (OAPT-2mm) - were calculated as follows:

$$Total\ dose_{pCT} = 10 \cdot d_{pCT} + 10 \cdot d_{(rCT1,pCT)} + 10 \cdot d_{(rCT2,pCT)}$$

where, d_{pCT} is the planned fraction dose on the *pCT*, $d_{(rCTn,pCT)}$ is the backpropagated fraction dose from the *rCTn* to the *pCT* (*n* taking values of 1 or 2), and $Total\ dose_{pCT}$ is the accumulated dose over all 30 fractions on the *pCT*.

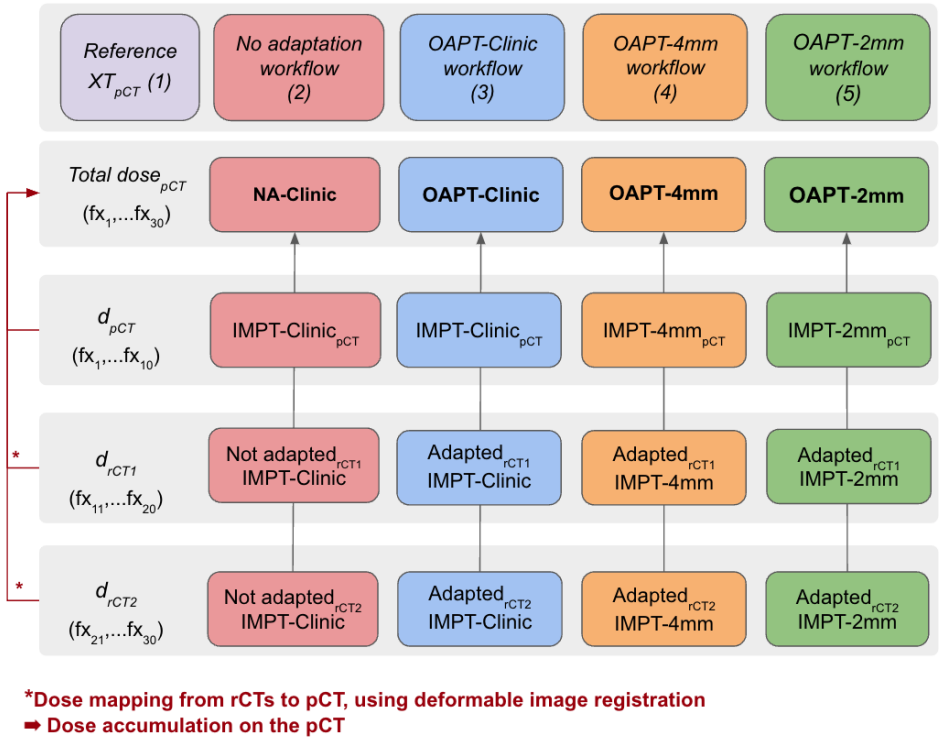


Figure 1. The total accumulated doses over all 30 fractions for 4 different plans: (1) the not adapted IMPT-Clinic (NA-Clinic) in red, adapted IMPT-Clinic (OAPT-Clinic) in blue, the adapted IMPT-4mm (OAPT-4mm) in orange and the adapted IMPT-2mm (OAPT-2mm) in green. Doses were back propagated using the deformable image registration between *pCT* and the corresponding *rCT*. The accumulated dose on the *pCT* ($Total\ dose_{pCT}$) was calculated through a weighted dose summation, accounting for 10 fractions for each the available CTs: 10 fractions for the *pCT* (1 to 10), 10 fractions for the *rCT1* (11 to 20) and 10 for the *rCT2* (21 to 30).

2.4. Normal Tissue Complication Probability (NTCP) metrics

NTCPs were calculated as described in the National Indication Protocol of Proton Therapy for Lung Carcinoma from The Netherlands for grade ≥ 2 radiation pneumonitis ($NTCP_{pneumonitis}$), grade ≥ 2 acute esophageal toxicity ($NTCP_{dysphagia}$) and 2 years mortality ($NTCP_{2ym}$) [31]. The detailed NTCP formulas and the clinical parameters involved, such as pulmonary co-morbidity, age, or gross tumor volume (GTV) size, can be found in the supplementary material SM3. The mean heart dose (MHD), mean lung dose (MLD), and mean esophageal dose (MED) were the dose metrics involved for $NTCP_{2ym}$, $NTCP_{pneumonitis}$, $NTCP_{dysphagia}$ and , respectively.

2.5. Evaluated scenarios

We consider for our calculations a simple but realistic setting of a single-room proton center with an 8-hour daily treatment capacity. For each NTCP model ($NTCP_{pneumonitis}$, $NTCP_{2ym}$ and $NTCP_{dysphagia}$), the NTCP values of the XT plans are considered as the reference. We investigate the variations on NTCP (ΔNTCP respect to the XT reference) in the following scenarios, also visually described in Figure 2:

- $\boldsymbol{S_{NA}}$: the no-adaptation (NA) scenario, where none of the patients are adapted and all of them undergo NA-Clinic. Therefore, if 14 patients are treated without plan adaptation in 8 hours (480 minutes), the base-line (BL) time per fraction per patient is 480/14 = 34.2 minutes.

$$\Delta NTCP_{S_{NA}} = (\sum_{p=1}^{p=14} NTCP_{XT} - \sum_{p=1}^{p=14} NTCP_{NA-Clinic})/14$$

- $\boldsymbol{S_{ideal}}$: the ideal scenario, where all patients are adapted following different $OAPT_x$ adaptation strategies: OAPT-Clinic (without setup error reduction), OAPT-4mm and OAPT-2mm. In this ideal scenario, the OAPT pipelines is considered instantaneous, no extra time per fraction is required with respect to no adaptation.

$$\Delta NTCP_{S_{ideal},OAPT_X} = (\sum_{p=1}^{p=14} NTCP_{XT} - \sum_{p=1}^{p=14} NTCP_{OAPT_X})/14$$

where $OAPT_X$, corresponds to OAPT-Clinic, OAPT-4mm and OAPT-2mm.

- **$S_{1,\dots,6}$**: A set of 6 scenarios in which, due to the increased time per fraction per patient experienced by a subset of OAPT-treated patients, not all the patients can be treated with protons. The remaining patients (from *i* = 1 to 6 patients) have to receive photon radiation therapy instead. In the scenario S_1, if the OAPT pipeline increases the time per fraction by 2.4 min (BL+2.4 = 36.9 min), only 13 patients could be treated with OAPT in the time span of 8h (13*36.9=480 min).

$$\Delta NTCP_{S(i=1,\dots,6),OAPT_X} = \left(\sum_{p=1}^{p=14} NTCP_{XT} - \left(\sum_{p=1}^{p=i} NTCP_{XT} + \sum_{p=i+1}^{p=14} NTCP_{OAPT_X}\right)\right)/14$$

In this formula the summatory term from p=1 to p=i, corresponds to the patients receiving XT instead of OAPT (in orange in Figure 2), while the term from p=i+1 to p=14 correspond to subset of OAPT-treated patients (in blue in Figure 2). In each $S_{1,\dots,6}$ the selected patients to receive XT instead of OAPT were those with a lower $\Delta NTCP_{total}$ = $NTCP_{XT}$- $NTCP_{OAPT\text{-}Clinic}$, being the $NTCP_{total}$ the probability of developping at least one of the side effects (see supplementary material SM4).

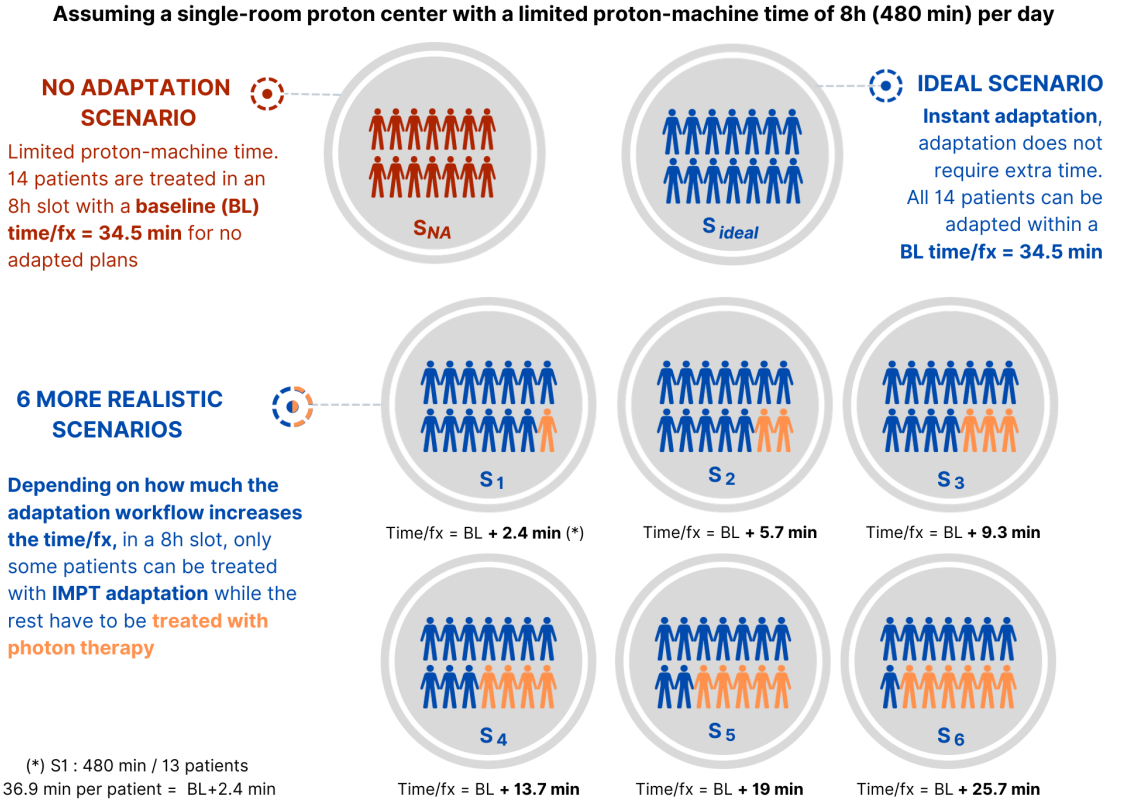


Figure 2. Summary of the scenarios in which the variation in normal tissue complication probabilities are evaluated. The no adaptation scenario (S_{NA}) corresponds to what would happen if the clinical IMPT plan (IMPT-Clinic) was delivered to the patients without adaptation (8h=480 min, 480min/14 patients = 34.5 min per fraction per patient). In $S_{ideal,}$ adaptation is considered instantaneous. On the contrary, in the more realistic scenarios ($S_{1,\dots,6}$) due to the increased time per fraction per patient experienced by a subset of OAPT-treated patients (blue patients) and the

limited proton-machine time of 480 min, not all the patients can be treated with protons. The remaining patients (orange patients) have to receive photon radiation therapy instead. For instance in S_1, only 13 patients can be treated with OAPT takes 2.4 min longer (*). Abbreviations: BL - Baseline, fx - fraction, NA - No adaptation, OAPT - Online adaptive proton therapy.

3. Results

In Figure 3, the mean variations in NTCP over the 14 patients provided by the OAPT strategies with respect to XT are represented for three possible treatment complications: (A) 2-year mortality, (B) grade ≥ 2 acute esophageal toxicity and (C) grade ≥ 2 radiation pneumonitis. As represented by the red dashed line in Figure 3., even if all patients are treated with the not adapted IMPT clinical plan (NA-Clinic) instead of using XT, this change of modality improves $NTCP_{2ym}$, $NTCP_{dysphagia}$ and $NTCP_{pneumonitis}$ by 6.9%, 6.1%, and 7.7%, respectively.

In the case of the ideal scenario (S_{ideal}) , in which all 14 patients can be instantaneously adapted with OAPT, we observed an improvement in mean ΔNTCP which varies according to the OAPT-strategy. For the 2-year-mortality side effect ($NTCP_{2ym}$), there is an improvement of 7.8% with respect to XT with OAPT-Clinic. For OAPT-2mm the gain in $\Delta NTCP_{2ym}$ goes up to 10.7%. A similar tendency was observed for $NTCP_{dysphagia}$ and $NTCP_{pneumonitis}$ in S_{ideal}. Compared to XT, NTCP improvements range from 6.1% to 13.6% and from 7.7% to 12.4%, for $NTCP_{dysphagia}$ and $NTCP_{pneumonitis}$, respectively.

When analyzing the intermediate scenarios ($S_{1,...,6}$) in Figure 3, a noticeable trend emerges: for all three adverse effects, the NTCP benefit relative to XT diminishes as the time required per fraction for OAPT increases. For both $NTCP_{2ym}$ and $NTCP_{pneumonitis}$ and without considering margin reduction (OAPT-Clinic), treatment fractions extended by more than 6 minutes do not imply an overall benefit in NTCP (S_2).

When applying setup error reduction, higher mean ΔNTCPs can be observed for longer fractions. For instance, for $NTCP_{2ym}$, OAPT-4mm provides an improvement of 8.0% respect to XT for S_3(+9.3 min) which is still higher than the 6.9% provided by NA-Clinic. However, for S_4(+13.7 min) the dosimetric benefits of the setup error reduction get diluted and only a slight benefit of 0.1% increase with respect to NA-Clinic (up to 7%) is obtained. For the most important setup error reduction (OAPT-2mm), $\Delta NTCP_{2ym}$ stops being significant respect to NA-Clinc (S_{NA}) when achieving scenario S_5(+19 min).

For $\Delta NTCP_{pneumonistis,}$ it is crucial to have a fast adaptation. At scenario S_4(+13.7 min), at approximately 15 extra minutes to perform adaptation, there is already almost no benefit for OAPT-2mm (7.8%) compared to NA-Clinic (7.7%). For this same side effect, if the

adaptation strategy does not include setup error reduction and adaptation is performed using the same setup errors as in the planning strategy (OAPT-Clinic), the benefit of OAPT quickly vanishes. All scenarios taking more than 2 extra minutes ($S_{2,3,4,5,6}$) have a lower average $\Delta NTCP_{pneumonistis}$ than the no adaptation scenario (S_{NA}).

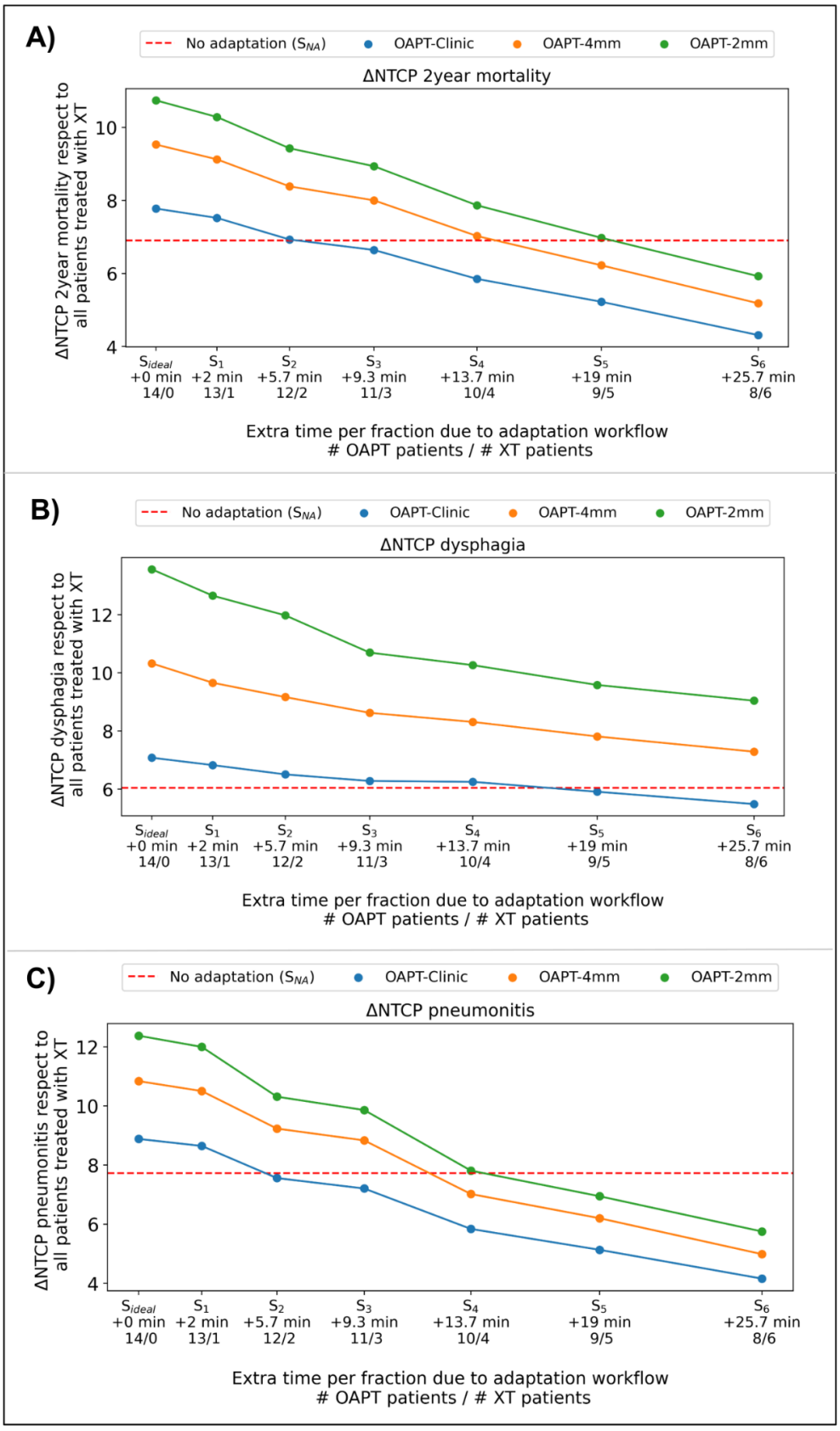


Figure 3. Variation in NTCP provided by three online adaptive proton therapy (OAPT) strategies for three complications: (A) 2-year mortality, (B) grade ≥ 2 acute esophageal toxicity, and (C) grade ≥ 2 radiation pneumonitis. The red line corresponds to the NTCP gain obtained from passing from XT

to not adapted IMPT. The points above this limit represent the situations in which the implementation of OAPT strategies provide a gain on mean NTCP considering all patients. For 2-year-mortality, a maximum of 6 minutes extra per fraction per patient is allowed in order to maintain an overall NTCP gain without setup error reduction (blue line). If setup error reductions are applied, 13 and 19 minutes extra could still provide an overall NTCP gain (orange and green lines).

4. Discussion

Our investigation sought to shed light on the balance between the clinical advantages of OAPT and the constraints imposed by treatment capacity and time limitations. To do so, we proposed an approach to quantitatively assess the impact of increased time per fraction on three NTCPs for lung cancer patients undergoing OAPT.

In the ideal scenario with instantaneous adaptation for all patients, substantial reductions in NTCP were achieved. Remarkably, the reductions in NTCP from transitioning from XT to not adapted IMPT were equivalent to those to those obtained by reducing setup errors to 2 mm. While clinical guidelines on setup error reduction in OAPT settings are yet to be established, our study highlights a notable trend: the dual benefit of selecting IMPT over XT and further gains from setup error reduction. Essentially, without setup error reduction and(or) instantaneous adaptation, the benefits of OAPT are minimal compared to the inherent advantages of IMPT. Our results align with the ones reported by Lalonde et al. in a cohort of 10 head and neck cancer patients. They suggest that clinically meaningful NTCP reductions were only achieved when combining setup uncertainty reduction and adaptation to anatomical changes [32].

Of particular interest is the observation that our study appears to have identified a threshold for the additional adaptation time that we can afford to keep an NTCP gain over our 14 patients. For 2-year mortality and grade>2 pneumonitis NTCPs (Figures 3.A and 3.C) this threshold was below 6 min (without setup error reduction) and between 13.7 and 19 minutes when applying setup error reduction. Feasibility within this time span has been already demonstrated for photon therapy using the Ethos delivery system [33]. Although such fully integrated adaptive solutions are not available yet for OAPT, several efforts are being made in that direction. However, it is important to highlight that according to our study, an extra time per fraction of around 15 minutes would only be worth it when OAPT comes together with a setup error reduction. Beyond these thresholds, the overall benefit of OAPT in terms of reducing NTCP diminishes, highlighting the importance of efficient and fast adaptation workflows.

Despite these insights, it is essential to acknowledge the study's limitations and the underlying assumptions. Our study was conducted within the context of a single-room

proton center with an 8-hour daily treatment capacity assuming maximum 14 patients treated (thus minimum 34.2 minutes per patient on average).These results may not be fully generalizable to facilities with different resources or capabilities. Moreover, the selection of patients to transition to XT in $S_{1,\ldots,6}$ was based on minimizing NTCP but different approaches, such as prediction of patients benefiting the most from OAPT, could be used for decision-making. Additionally, our comparison considered the quality of adapted plans in an online setting to be equivalent to those generated offline, an assumption that requires further investigation due to the possible uncertainties related to automatic re-contouring without manual correction or approximations in the optimization algorithms in an online adaptive setting. Moreover, we operated under the assumption that the planned XT dose distributions were representatives of the actual delivered dose. However, it could be worth exploring the potential advantages of integrating online adaptive in the XT branch as well. Finally, we have used setup errors in robust optimisation according to the clinical margins used for the XT plans. Such assumption may be challenged by the fact that state-of-the-art clinical practice includes in general some form of offline adaptive procedure. The gain of using an online adaptive procedure may then be somewhat diminished. However, the setup error reduction that can be achieved in an offline adaptive procedure is not clear to us, as it will depend at least on practical considerations about the adaptation rate. Additionally, offline adaptive strategies will not be fully immunized against inter-fraction motion; therefore, substantial margins will probably have to be used.

Furthermore, a noteworthy aspect of this study is the use of dosimetric values from accumulated dose distributions to compute NTCPs. The dose mapping using deformable image registration might introduce uncertainties, especially in regions with a high dose gradient and in the presence of volume changes, such as tumor shrinkage [34]. However, the NTCP calculations are based on metrics of average doses to organs at risk that do not present critical volume changes, such as the heart, the lungs, or the esophagus (see supplementary material Table SM1.2). Therefore, although dose accumulation could affect the accuracy of NTCP calculations we believe there is a limited impact, as also demonstrated by Nennof et al. [35], and that this approach still provides valuable insights.

In conclusion, this study highlights the interplay between treatment adaptation and time constraints in the context of online adaptive proton therapy. The discernible improvements in NTCP underscore the potential clinical advantages of OAPT, particularly in cases with heightened risks of treatment-related toxicities, such as 2-year mortality. However, these benefits must be balanced against the practical limitations of treatment capacity and the extra time per fraction required for adaptation. It is essential for the field of proton therapy to recognize this balance and refine the adaptation process to optimize patient outcomes effectively.

Acknowledgements

Edmond Sterpin is involved in collaborations with Ion Beam Applications and RaySearch laboratories.

References

1. Nierer L, Eze C, da Silva Mendes V, Braun J, Thum P, von Bestenbostel R, et al. Dosimetric benefit of MR-guided online adaptive radiotherapy in different tumor entities: liver, lung, abdominal lymph nodes, pancreas and prostate. Radiat Oncol. 2022;17:53.

2. Byrne M, Archibald-Heeren B, Hu Y, Teh A, Beserminji R, Cai E, et al. Varian ethos online adaptive radiotherapy for prostate cancer: Early results of contouring accuracy, treatment plan quality, and treatment time. J Appl Clin Med Phys. 2022;23:e13479.

3. Yoon SW, Lin H, Alonso-Basanta M, Anderson N, Apinorasethkul O, Cooper K, et al. Initial Evaluation of a Novel Cone-Beam CT-Based Semi-Automated Online Adaptive Radiotherapy System for Head and Neck Cancer Treatment - A Timing and Automation Quality Study. Cureus. 2020;12:e9660.

4. El-Bared N, Portelance L, Spieler BO, Kwon D, Padgett KR, Brown KM, et al. Dosimetric Benefits and Practical Pitfalls of Daily Online Adaptive MRI-Guided Stereotactic Radiation Therapy for Pancreatic Cancer. Pract Radiat Oncol. 2019;9:e46–54.

5. Christiansen RL, Dysager L, Hansen CR, Jensen HR, Schytte T, Nyborg CJ, et al. Online adaptive radiotherapy potentially reduces toxicity for high-risk prostate cancer treatment. Radiother Oncol. 2022;167:165–71.

6. Sibolt P, Andersson LM, Calmels L, Sjöström D, Bjelkengren U, Geertsen P, et al. Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region. Phys Imaging Radiat Oncol. 2021;17:1–7.

7. Henke LE, Olsen JR, Contreras JA, Curcuru A, DeWees TA, Green OL, et al. Stereotactic MR-Guided Online Adaptive Radiation Therapy (SMART) for Ultracentral Thorax Malignancies: Results of a Phase 1 Trial. Adv Radiat Oncol. 2019;4:201–9.

8. Hoegen P, Lang C, Akbaba S, Häring P, Splinter M, Miltner A, et al. Cone-Beam-CT Guided Adaptive Radiotherapy for Locally Advanced Non-small Cell Lung Cancer Enables Quality Assurance and Superior Sparing of Healthy Lung. Front Oncol. 2020;10:564857.

9. Trnkova P, Zhang Y, Toshito T, Heijmen B, Richter C, Aznar MC, et al. A survey of practice patterns for adaptive particle therapy for interfractional changes. Phys Imaging Radiat Oncol. 2023;26:100442.

10. Ribeiro CO, Visser S, Korevaar EW, Sijtsema NM, Anakotta RM, Dieters M, et al. Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion: Robust optimised plan evaluation by means of patient and machine specific information. Radiother Oncol. 2021;157:210–8.

11. Szeto YZ, Witte MG, van Kranen SR, Sonke J-J, Belderbos J, van Herk M. Effects of anatomical changes on pencil beam scanning proton plans in locally advanced NSCLC patients. Radiother Oncol. 2016;120:286–92.

12. Hoffmann L, Alber M, Jensen MF, Holt MI, Møller DS. Adaptation is mandatory for intensity modulated proton therapy of advanced lung cancer to ensure target coverage. Radiother Oncol. 2017;122:400–5.

13. Bernatowicz K, Geets X, Barragan A, Janssens G, Souris K, Sterpin E. Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration. Phys Med Biol. 2018;63:085018.

14. Villarroel EB, Geets X, Sterpin E. Online adaptive dose restoration in intensity modulated proton therapy of lung cancer to account for inter-fractional density changes. Physics and Imaging in Radiation Oncology. 2020;15:30–7.

15. McComas KN, Yock A, Darrow K, Shinohara ET. Online Adaptive Radiation Therapy and Opportunity Cost. Adv Radiat Oncol. 2023;8:101034.

16. Verma V, Mishra MV, Mehta MP. A systematic review of the cost and cost-effectiveness studies of proton radiotherapy. Cancer. 2016;122:1483–501.

17. Huang D, Frank SJ, Verma V, Thaker NG, Brooks ED, Palmer MB, et al. Cost-Effectiveness Models of Proton Therapy for Head and Neck: Evaluating Quality and Methods to Date. Int J Part Ther. 2021;8:339–53.

18. Langendijk JA, Hoebers FJP, de Jong MA, Doornaert P, Terhaard CHJ, Steenbakkers RJHM, et al. National Protocol for Model-Based Selection for Proton Therapy in Head and Neck Cancer. Int J Part Ther. 2021;8:354–65.

19. Draguet C, Barragán-Montero AM, Vera MC, Thomas M, Populaire P, Defraene G, et al. Automated clinical decision support system with deep learning dose prediction and NTCP models to evaluate treatment complications in patients with esophageal cancer. Radiother Oncol. 2022;176:101–7.

20. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2000;47:1121–35.

21. Errors and margins in radiotherapy. Semin Radiat Oncol. 2004;14:52–64.

22. Cummings D, Tang S, Ichter W, Wang P, Sturgeon JD, Lee AK, et al. Four-dimensional Plan Optimization for the Treatment of Lung Tumors Using Pencil-beam Scanning Proton Radiotherapy. Cureus. 2018;10:e3192.

23. Albertini F, Hug EB, Lomax AJ. Is it necessary to plan with safety margins for actively scanned proton therapy? Phys Med Biol. 2011;56:4399–413.

24. van der Voort S, van de Water S, Perkó Z, Heijmen B, Lathouwers D, Hoogeman M. Robustness Recipes for Minimax Robust Optimization in Intensity Modulated Proton Therapy for Oropharyngeal Cancer Patients. Int J Radiat Oncol Biol Phys. 2016;95:163–70.

25. Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, et al.

Practical robustness evaluation in radiotherapy - A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol. 2019;141:267–74.

26. Badiu V, Souris K, Buti G, Villarroel EB, Lambrecht M, Sterpin E. Improved healthy tissue sparing in proton therapy of lung tumors using statistically sound robust optimization and evaluation. Phys Med. 2022;96:62–9.

27. Wanet M, Sterpin E, Janssens G, Delor A, Lee JA, Geets X. Validation of the mid-position strategy for lung tumors in helical TomoTherapy. Radiother Oncol. 2014;110:529–37.

28. Dosimetric Impact of Simulated Daily Adaptive Radiotherapy with Significantly Reduced Setup Margins in the Definitive Treatment of Head and Neck Cancer. International Journal of Radiation Oncology*Biology*Physics. 2022;114:e590.

29. Byrne M, Teh AYM, Archibald-Heeren B, Hu Y, Rijken J, Luo S, et al. Intra-fraction motion and margin assessment for Ethos online adaptive radiotherapy treatments of the prostate and seminal vesicles. Adv Radiat Oncol. 2023;101405.

30. Mercieca S, Belderbos JSA, van Herk M. Challenges in the target volume definition of lung cancer radiotherapy. Transl Lung Cancer Res. 2021;10:1983–98.

31. Landelijk Platform voor Radiotherapie bij Longtumoren (LPRL) & Landelijk Platform Protonentherapie (LPPT). Landelijk Indicatie Protocol Protonen Therapie Longcarcinoom [Internet]. 2019 [cited 2023 Aug 17]. Available from: https://nvro.nl/images/documenten/rapporten/LIPP_longen_final_01122019.pdf

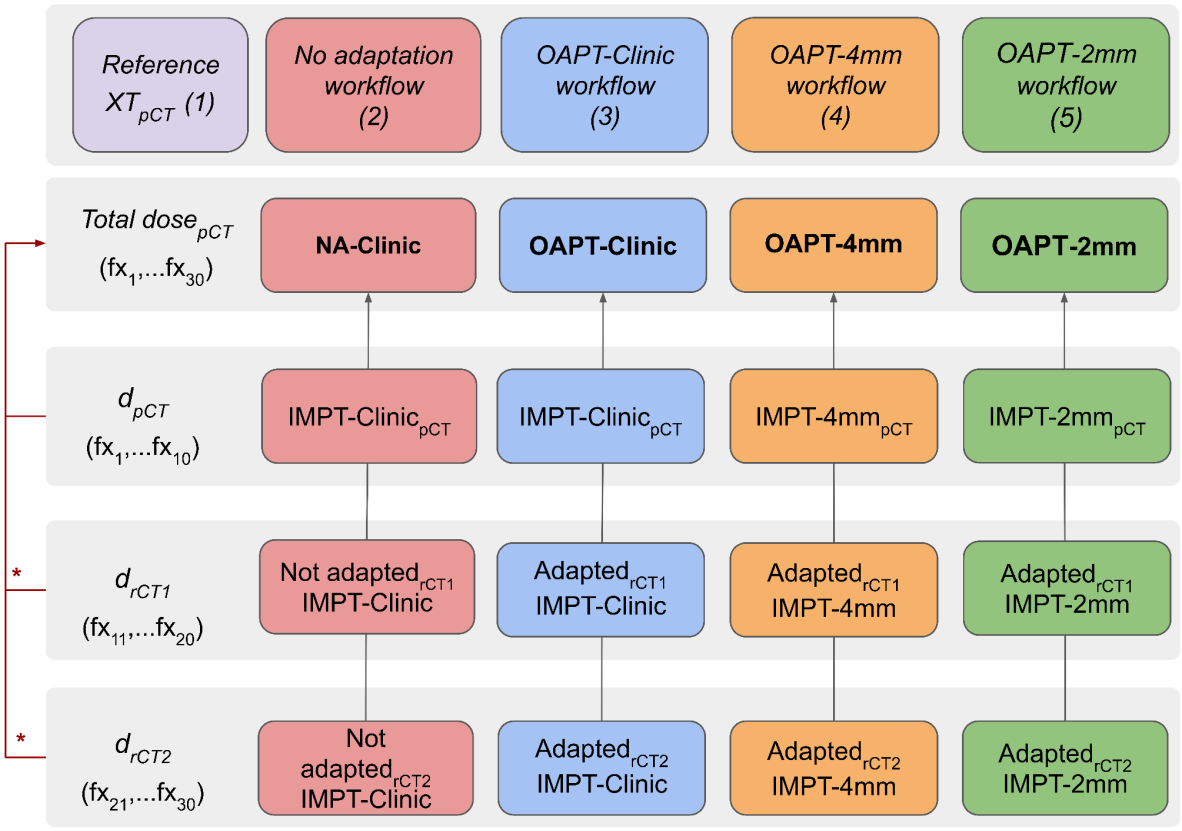
32. Lalonde A, Bobić M, Sharp GC, Chamseddine I, Winey B, Paganetti H. Evaluating the effect of setup uncertainty reduction and adaptation to geometric changes on normal tissue complication probability using online adaptive head and neck intensity modulated proton therapy. Phys Med Biol [Internet]. 2023;68. Available from: http://dx.doi.org/10.1088/1361-6560/acd433

33. Stanley DN, Harms J, Pogue JA, Belliveau J-G, Marcrom SR, McDonald AM, et al. A roadmap for implementation of kV-CBCT online adaptive radiation therapy and initial first year experiences. J Appl Clin Med Phys. 2023;24:e13961.

34. Murr M, Brock KK, Fusella M, Hardcastle N, Hussein M, Jameson MG, et al. Applicability and usage of dose mapping/accumulation in radiotherapy. Radiother Oncol. 2023;182:109527.

35. TCP and NTCP Calculations Based on Treatment Doses Instead of Planned Doses for Daily Adaptive Proton Therapy of Lung Cancer. International Journal of Radiation Oncology*Biology*Physics. 2021;111:e133.

Figures



***Dose mapping from rCTs to pCT, using deformable image registration**
➡ Dose accumulation on the pCT

Figure 1. The total accumulated doses over all 30 fractions for 4 different plans: (1) the not adapted IMPT-Clinic (NA-Clinic) in red, adapted IMPT-Clinic (OAPT-Clinic) in blue, the adapted IMPT-4mm (OAPT-4mm) in orange and the adapted IMPT-2mm (OAPT-2mm) in green. Doses were back propagated using the deformable image registration between *pCT* and the corresponding *rCT*. The accumulated dose on the *pCT* (*Total dose*$_{pCT}$) was calculated through a weighted dose summation, accounting for 10 fractions for each the available CTs: 10 fractions for the *pCT* (1 to 10), 10 fractions for the *rCT1* (11 to 20) and 10 for the *rCT2* (21 to 30).

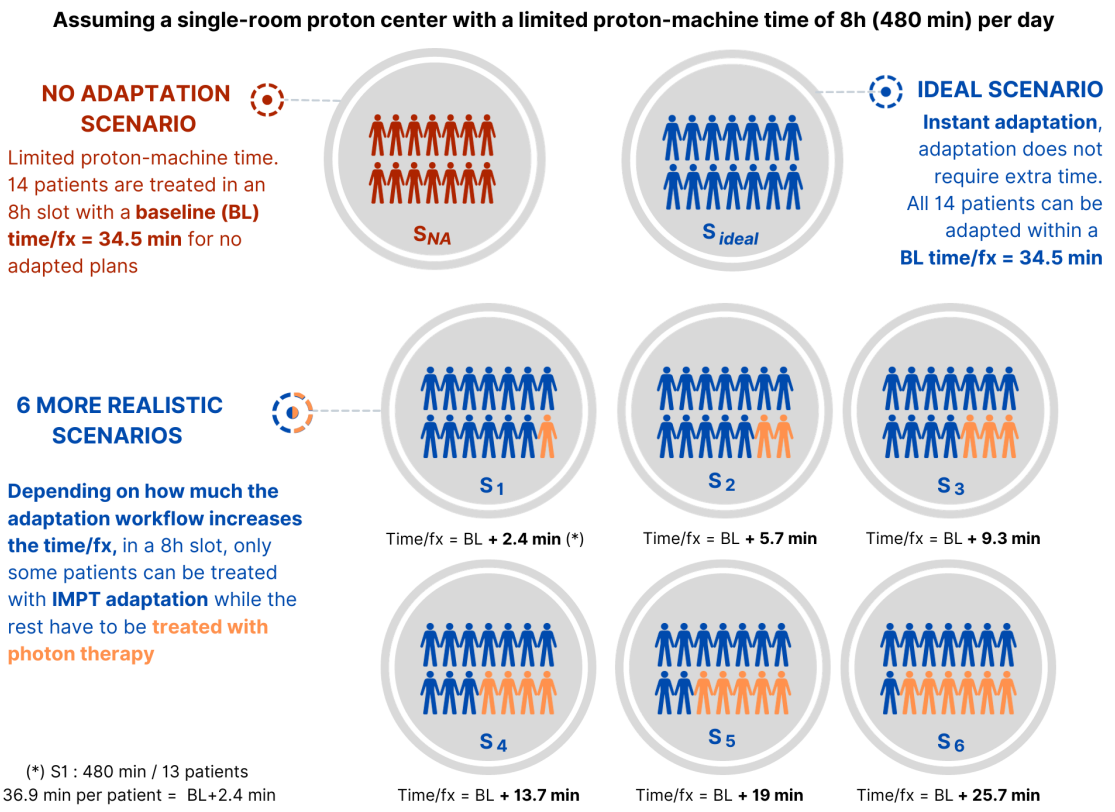


Figure 2. Summary of the scenarios in which the variation in normal tissue complication probabilities are evaluated. The no adaptation scenario (S_{NA}) corresponds to what would happen if the clinical IMPT plan (IMPT-Clinic) was delivered to the patients without adaptation (8h=480 min, 480min/14 patients = 34.5 min per fraction per patient). In $S_{ideal,}$ adaptation is considered instantaneous. On the contrary, in the more realistic scenarios ($S_{1,...,6}$) due to the increased time per fraction per patient experienced by a subset of OAPT-treated patients (blue patients) and the limited proton-machine time of 480 min, not all the patients can be treated with protons. The remaining patients (orange patients) have to receive photon radiation therapy instead. For instance in S_1, only 13 patients can be treated with OAPT takes 2.4 min longer (*). Abbreviations: BL - Baseline, fx - fraction, NA - No adaptation, OAPT - Online adaptive proton therapy.

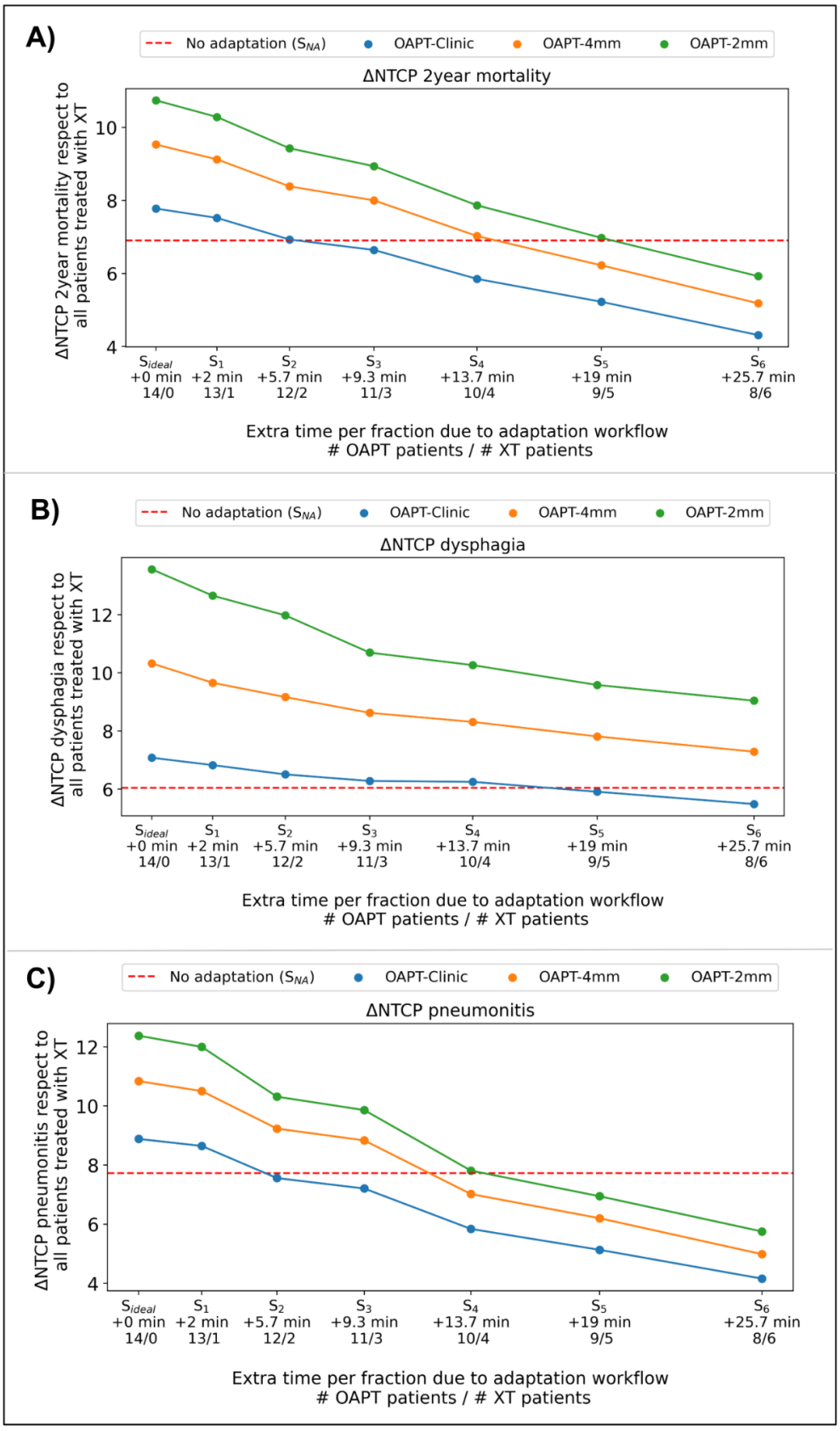


Figure 3. Variation in NTCP provided by three online adaptive proton therapy (OAPT) strategies for three complications: (A) 2-year mortality, (B) grade ≥ 2 acute esophageal toxicity, and (C) grade ≥ 2 radiation pneumonitis. The red line corresponds to the NTCP gain obtained from passing from XT

to not adapted IMPT. The points above this limit represent the situations in which the implementation of OAPT strategies provide a gain on mean NTCP considering all patients. For 2-year-mortality, a maximum of 6 minutes extra per fraction per patient is allowed in order to maintain an overall NTCP gain without setup error reduction (blue line). If setup error reductions are applied, 13 and 19 minutes extra could still provide an overall NTCP gain (orange and green lines).