

1 **Evaluating the adaptation rates for esophageal cancer: impact of the**
2 **setup error contribution and of the dosimetric threshold**

3 Authors: Camille Draguet^{1,2,*}, Pieter Populaire^{2,3}, Macarena Chocan Vera¹,
4 Karin Haustermans^{2,3}, John A. Lee¹, Ana M. Barragán-Montero¹ and Ed-
5 mond Sterpin^{1,2}.

6 ¹ UCLouvain, Institut de Recherche Expérimentale et Clinique, Molecular
7 Imaging Radiotherapy and Oncology (MIRO), 1200 Brussels, Belgium;

8 ² KU Leuven, Department of Oncology, Laboratory of Experimental Radio-
9 therapy, 3000 Leuven, Belgium;

10 ³ University Hospitals Leuven, Department of Radiation Oncology, 3000 Leu-
11 ven, Belgium;

12 **corresponding author:*

13 camille.draguet@uclouvain.be

14 Avenue Hippocrate 55/B1.55.01, B-1200 Bruxelles, Belgium

15 **Introduction**

16 Locally advanced esophageal cancer (EC) is commonly treated using neoad-
17 juvant chemo-radiotherapy followed by surgery. The standard radiother-
18 apy delivery technique for EC is intensity-modulated radiotherapy (IMRT).
19 However, proton therapy (PT) is recognized for its ability to produce a more
20 conformal dose distribution. It has been shown to reduce toxicity by better
21 sparing organs-at-risk (OARs), such as the lungs and heart [1–3]. In this con-
22 text, the PROTECT clinical trial has been established to evaluate the benefits
23 of PT compared to photon-based radiotherapy (XT) for EC [4], focusing on
24 postoperative pulmonary complications within 0-90 days after surgery and
25 the need for adaptive treatment in both modalities.

26 The XT or PT regimens are typically delivered in smaller fraction doses
27 over several days. The treatment delivery is, therefore, prone to uncertain-
28 ties related to the patient’s anatomy, which often changes during treatment.
29 The aim of adaptive radiotherapy (ART) is to address these uncertainties
30 and ensure precise treatment delivery by acquiring longitudinal images (e.g.,
31 weekly or daily) and adjusting the treatment plan according to the updated
32 anatomy [5]. ART has demonstrated the potential to improve patient out-

33 comes [6–8], but its implementation is known to be resource-intensive. As a
34 result, triggered adaptation is frequently employed, which refers to adapting
35 the treatment plan only if a specific threshold is exceeded [5].

36 Usually, the criteria for adaptation are qualitative and rely on the visual
37 inspection of the planned dose distribution, transposed onto the anatomy of
38 the day. Kwint et al. [9] proposed a semi-quantitative protocol to prioritize
39 anatomical changes in lung cancer radiotherapy, primarily based on the posi-
40 tion of the gross tumor volume (GTV) relative to the planning target volume
41 (PTV). Möller et al. [10] established a set of criteria based on the evaluation of
42 the positions of lymph nodes and organs-at-risk (OARs), with dose restora-
43 tion triggered if these criteria are exceeded. However, to eliminate the bur-
44 den of visually inspecting each fraction and reduce inter-observer variability,
45 adaptation should ideally be triggered by dosimetric criteria. The benefit of
46 adaptation can then be evaluated by accumulating the doses recalculated on
47 repeated images, which provides a reliable estimate of the delivered dose.

In radiotherapy, geometrical uncertainties related to patient positioning and beam delivery are typically handled either through margins (most commonly in XT) or through robust optimization (most commonly in PT) [11, 12]. In XT, these uncertainties are generally incorporated by expanding the Clinical Target Volume (CTV) into a Planning Target Volume (PTV) using an appropriate margin. In contrast, robust optimization tackles these uncertainties directly within the optimization process, by simulating scenarios corresponding to different realizations of the uncertainties and computing the plan parameters according to the worst of all scenarios. In PT, because of the finite depth-dose profile of protons, the “static dose cloud approximation” does not hold, meaning that geometric shifts do not translate predictably to dose shifts and making margins insufficient; robust optimization is therefore often used instead of margins.

When evaluating the recalculated dose on the new anatomy, addressing its robustness against treatment delivery errors is essential, particularly for PT plans. For both XT and PT, robustness evaluation typically involves applying rigid shifts of the patient (setup errors), and, in the case of PT, additionally scaling of the densities to simulate proton range errors, thereby resulting in different error scenarios. Commonly, the treatment plan’s robustness is assessed against either the worst of all scenarios relative to a specific dosimetric objective function (i.e., the worst-case scenario) or an aggregation of the voxelwise worst-case dose across all scenarios (i.e., the voxelwise worst dose in this study) [13]. For the target volume, the voxelwise worst dose includes voxels receiving the lowest dose across all scenarios (i.e., the voxelwise minimum dose), while in the rest of the body, it includes voxels receiving the highest dose among all scenarios (i.e., the voxelwise maximum dose).

The setup error considered in the robust evaluation aggregates uncertainties of different types, including the actual uncertainties related to patient setup and inter- and intra-fraction motion, excluding breathing motion. The latter can be captured with a 4D-CT. The use of daily imaging allows to correctly position the patient before each fraction and partly reduce this aggregated setup error, which justifies relying on a lower maximum setup error when evaluating the dose recalculated on the new anatomy [14–16], referred in the following as the residual setup error [17]. However, there are no established guidelines or models for the appropriate selection of these residual setup errors. It can be argued that a newly acquired CT scan only partially

reduces uncertainties, as intra-fractional motion (excluding breathing) is not fully captured in the image [18, 19]. Moreover, delineation errors and associated unpredicted isocenter shifts must also be accounted for. The choice of residual setup error will impact the adaptation rate and must be carefully determined.

This study investigates adaptive strategies for XT and PT and evaluates the impact of triggering dosimetric criteria and selected setup error on the adaptation frequency for EC. A brief literature review was conducted to compile existing dosimetric adaptation triggering criteria and to compare them with our adaptation protocol, which is aligned with the PROTECT guidelines. The influence of the evaluation setup errors is also reported. Additionally, we compared the adapted and non-adapted doses in terms of target coverage and OAR dose, showing the benefit of adaptation for our EC patient population.

Materials and Methods

Literature review

Public databases such as PubMed and Google Scholar were searched using keywords that combined terms like "Adaptive radiotherapy", "Adaptive proton therapy", "Replanning protocol", "Triggering adaptation", and "Esophageal cancer". We focused on studies that specifically referenced dosimetric criteria for triggering adaptation in patients with EC undergoing XT or PT. The dose distributions recalculated on repeated CT (rCT) scans were evaluated against these dosimetric criteria to determine if treatment adaptation was warranted. This search resulted in five relevant papers outlining adaptation protocols for EC, with a summary of key protocol features provided in Table 1.

Nyeng et al. [20] investigated the dosimetric consequences of anatomical changes occurring during XT treatment in 29 EC patients to establish criteria for a decision-support system tool that assists clinicians in determining the need for adaptation. Anatomical changes were considered significant if they resulted in a decrease in CTV V95% of 1% on the mid-ventilation phase of a 4D-CT taken one week after treatment initiation, compared to the initial

117 plan, and if the PTV V95% decreased by 3%. The plan was adapted if these
118 changes persisted on the CBCTs of the following three fractions. These cri-
119 teria are inspired by the work of Möller et al. [10], which explored the dosi-
120 metric benefits of plan adaptation in lung IMRT adaptation.

121 Shamshad et al. [21] explored the impact of patient positioning based on
122 either bone or soft-tissue matching on PT plan adaptation for 43 EC patients.
123 Adaptation was triggered if the iCTV V95% (aggregation of the CTV on all
124 the 4D-CT phases) was below 98% on the nominal dose and below 96% on
125 the worst-case dose, evaluated with a 2mm residual setup and a $\leq 3\%$ density
126 variation. Additionally, adaptation could be triggered if, on any respiratory
127 phase, the CTV V95% was less than 96%. These criteria align with the PRO-
128 TECT adaptation guidelines recommended for assessing the adaptation need
129 in the ten first fractions.

130 In the study by Visser et al. [22], the authors evaluated the robustness
131 of adaptive photon and proton radiotherapy treatment plans for 20 EC pa-
132 tients. Three radiation oncologists rated the dose distributions when the
133 voxelwise minimum iCTV D98% was below 96% and determined whether
134 plan adaptation was necessary. The authors reported that adaptation was
135 generally required when the voxelwise minimum iCTV D98% dropped below
136 94%. Anakotta et al. [15] applied the same criteria to evaluate iCTV cover-
137 age in recalculated VMAT and IMPT plans for 19 EC patients. Additionally,
138 they ensured Spinal Canal D0.1cm³ < 51 Gy, which would otherwise trigger
139 adaptation. This last criterion is not considered in this study, but was never
140 exceeded in our patient cohort.

141 Canters et al. [23] investigated the robustness of EC IMPT treatment
142 plans against anatomical variations on 60 patients. Target coverage (iCTV
143 V94% < 98%) was assessed on the voxelwise minimum dose, recalculated
144 on weekly repeated 4D-CT scans with a 2mm residual setup error and a 3%
145 range error. This criterion is equivalent to that used by Visser et al. [22]
146 and Anakotta et al. [15]. The primary reason for adaptation was reported as
147 target underdosage, with additional factors including initial plan improve-
148 ments, increased dose to OARs, and significant delineation errors.

149 In addition to the papers listed in Table 1, we can also mention the work
150 of Ma et al. in lung ART [24]. The authors compared two complete sets
151 of criteria, one based on target coverage and OAR dose observed in the re-
152 calculated dose distribution and the other on the degradation of the initial

153 plan quality. These two different sets resulted in varying adaptation rates.
154 Veiga et al. [25] also used a combination of these two types of criteria (initial
155 plan quality degradation versus OAR dose evaluation) for their adaptation
156 protocol in lung IMPT. In head and neck cancers, implemented dosimetric
157 protocols are primarily triggered in case of target underdosage [26, 27]. The
158 recent study by Gros et al. [27] further demonstrated how their clinically
159 adapted decision-support software can help predict treatment courses re-
160 quiring adaptation four fractions in advance.

161 The criteria presented in Table 1 were compared with our adaptation
162 protocol, following the PROTECT guidelines. Details about our patient co-
163 hort, planning strategies, and adaptation criteria are given in the following
164 sections.

Table 1: Existing dosimetric criteria, applied to trigger plan adaptation for EC patients. iCTV and iPTV correspond to the aggregation of CTV and PTV over all 4D-CT phases, respectively.

Authors	Evaluation CT	Rigid Registration	Adaptation Criteria	Modality	Adaptation Rate (%)	Robust Evaluation Parameters
Nyeng et al. [20] 2015	One additional 4D-CT, evaluation on the mid-ventilation phase, 10 days after the start of treatment	Bone-match	Decrease in CTV V95% > 1% and in PTV V95% > 3%, compared to the initial plan. Spinal Canal Dmax > 45 Gy. Adaptation occurred if anatomical changes persisted over 3 fractions.	XT (not specified)	13/29 (44.8%) based on dosimetric criteria only; 9/29 (31%) after observation on 3 fractions' CBCTs.	None
Anakotta et al. [15] 2020	Four additional weekly 4D-CT	Bone-match	iCTV D98% < 94% Spinal Canal D0.1cm ³ > 51 Gy	VMAT IMPT	2/19 (10.5%) 4/19 (21%)	2mm setup error 0% range error 2mm setup error 3% range error
Samshad et al. [21] 2022	One additional 4D-CT taken at a median time of 10 [5-15] days after the start of treatment	Bone-match, soft-tissue match	iCTV V95% < 98% to the nominal dose distribution and iCTV V95% < 96% to the worst-case scenario	IMPT	5/43 (11.6%) for bone-match 0/43 (0%) for soft-tissue match	2mm setup error 3.5% range error
Visser et al. [22] 2022	Four additional weekly 4D-CT	Not specified	iCTV D98% < 96% on voxelwise minimum dose. This triggered coverage inspection by 3 radiation oncologists. Adaptation generally required if iCTV D98% < 94%.	VMAT IMPT	6/19 (31.5%), one of them required a second replanning at week 3. 9/19 (47.3%), 5 of them required a second replanning at week 2 (1 patient), week 3 (2), and week 4 (2).	2mm setup error 0% range error 2mm setup error 3% range error
Canterers et al. [23] 2024	Four or five additional weekly 4D-CT	Bone-match	iCTV V94% < 98%	IMPT	14/60 (23.3%), 8 of them due to target underdosage, 6 for other reasons	2mm setup error 3% range error

165 Patient cohort and planning strategies

166 Between 2016 and 2018, a cohort of 18 patients diagnosed with locally ad-
167 vanced EC underwent neoadjuvant chemoradiation therapy followed by sur-
168 gery at the University Hospital of Leuven (UZ Leuven). The collection and
169 analysis of patient data for this study were granted ethical approval by the In-
170 stitutional Ethical Review Board of UZLeuven (S59667). To ensure database
171 consistency and quality, all patients were systematically planned for IMRT
172 and proton beam scanning (PBS), adhering to the treatment planning guide-
173 lines outlined in PROTECT [4]. Specific beam configurations are given in
174 Supplementary Materials, Table S1. Guidelines for target delineation can
175 be found in the proposal of Thomas et al. [28]. The prescribed treatment
176 regimen involved a total dose of 50.4 Gy delivered over 28 fractions. See
177 Supplementary Materials for more information.

178 The 18 individuals underwent a repeated CT scan (rCT) after approxi-
179 mately three weeks of treatment (~ 15 fractions). These repeated CT scans
180 were acquired systematically, without any a priori selection of the patients.
181 Both the planning CT (pCT) and the rCT were acquired on a Siemens So-
182 matom Sensation Open and the corresponding average CT scans are used
183 for this study. In the following, the acronyms rCT and pCT refer to the av-
184 erage repeated CT and the average planning CT, respectively. Patient char-
185 acteristics, including tumor staging and location on the esophageal wall, are
186 reported in Supplementary Materials, Table SS4. The average rCT scans were
187 aligned to the average pCT scans using rigid registration algorithm with a
188 soft-tissue match, using the algorithm embedded in RayStation (translations
189 and rotations allowed), and new contours were delineated on each rCT. The
190 doses were then recomputed with the initial plan parameters on the average
191 rCT to obtain the “non-adapted” (NA) doses. These NA doses were com-
192 pared to the manually adapted doses (MA dose). In the manual branch (MA),
193 two plans were created per patient: an IMRT plan using a 7 mm iPTV, and
194 an IMPT PBS plan using robust optimization [29] on iCTV with 7 mm setup
195 error and 2.6% range uncertainty. A physician subsequently validated each
196 MA plan. iPTV and iCTV correspond to the aggregation of PTV and CTV
197 over all 4D-CT phases, respectively.

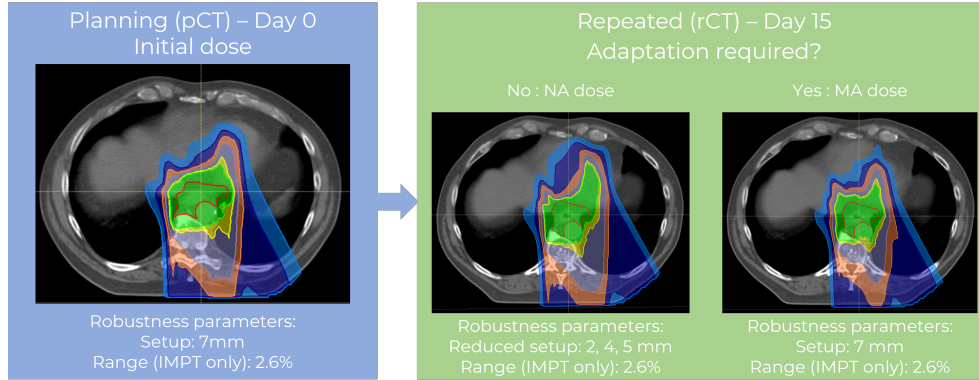


Figure 1: Workflow for treatment plan preparation. For each patient, a non-adapted plan (NA dose) and a manually adapted plan (MA dose) were created. The dose distributions shown correspond to typical IMPT doses, but the workflow applies to both IMRT and IMPT. The robustness evaluation parameters associated with each plan are indicated below each dose (setup parameters for IMRT; setup and range parameters for IMPT).

Adaptation rate

Impact of the adaptation criteria

In the PROTECT guidelines, adaptation is indicated when clinical goals for the nominal NA dose to the iCTV, spinal cord, and body are not met (without considering any tolerance), as outlined in Table 2. The target coverage is also evaluated on the worst-case scenario dose. These results were then compared to the criteria presented in Table 1. Since the adaptation protocol proposed by Shamshad et al. [21] is also based on the PROTECT guidelines (see Supplementary Materials, Table S5) but implemented at a different time step, we did not report adaptation rates specific to their criteria. Notably, we did not have access to phase images from our rCTs. Therefore, we hypothesized that the criteria established by Nyeng et al. [20], applied initially to the mid-ventilation phase, could be similarly assessed on the average rCT.

Table 2: PROTECT adaptation criteria, from week 3 to week 4

Fraction 11-21	
iCTV	
<i>Nominal</i>	$V95\% > 97\%$
<i>All scenarios</i>	$V95\% > 95\%$
Spinal canal	
<i>Nominal</i>	$D0.5\text{cm}^3 < 50 \text{ Gy}$
Body	
<i>Nominal</i>	$D2\text{cm}^3 < 110\%$
	$D7\text{cm}^3 < 107\%$

211 The comparison between the different dosimetric criteria led to differing
 212 adaptation rates. Criteria related to iCTV target coverage are illustrated on
 213 a representative iCTV DVH curve in Figure 2, highlighting the differences
 214 between them. However their implications in terms of dosimetric require-
 215 ments are similar.

216 Impact of the residual setup error

217 Our literature review revealed that authors generally consider a residual
 218 setup error of 2 mm when evaluating the NA dose [15, 21, 22], even though
 219 this choice is not supported by a consensus or by a clear procedure for de-
 220 termining this margin. This recommendation has also been adopted by the
 221 PROTECT protocol. However, as mentioned in the introduction, it is chal-
 222 lenging to justify the use of a specific setup margin due to the uncertainty
 223 in determining which portion of the intra-fraction residual setup error is
 224 already realized in the new image. This setup error, in combination with
 225 contouring inaccuracies and associated isocenter positioning, can only be
 226 addressed by selecting an appropriate setup margin. Various strategies were
 227 explored, incorporating 2, 4, and 5 mm residual setup errors. The corre-
 228 sponding adaptation rates are reported.

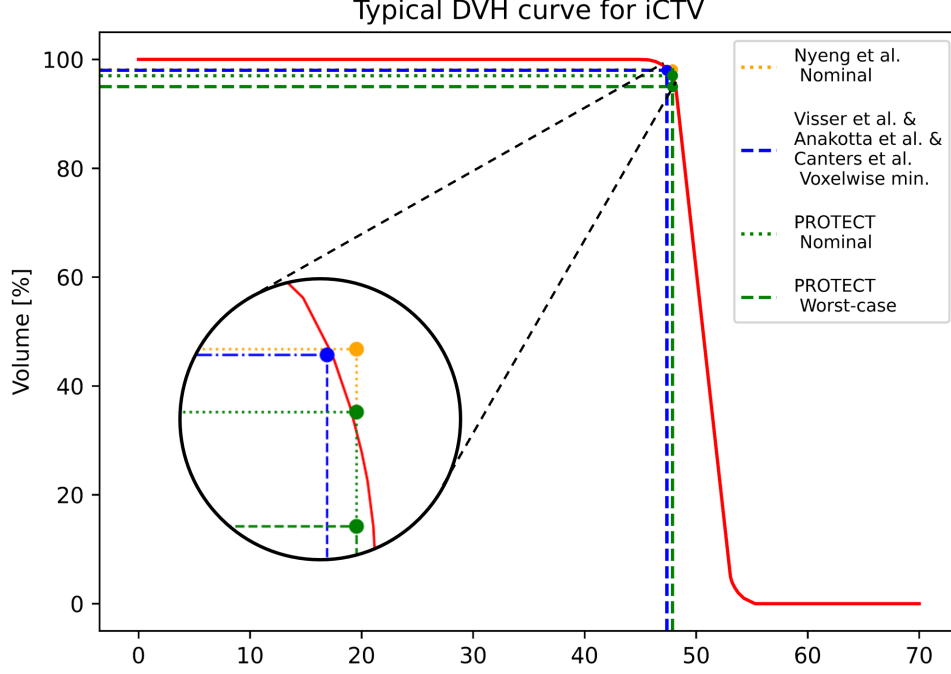


Figure 2: The criteria for adaptation are illustrated on a typical iCTV DVH curve. For the criterion used by Nyeng et al. [20], which is typically patient-specific, we show the 1% decrease in iCTV V95% compared to the median initial iCTV V95% calculated for our patient cohort. Both nominal and worst-case/voxelwise minimum criteria are illustrated on the same curve for ease of representation but should be considered separately.

229 Benefits of adaptation for EC

230 Additionally, the no-adaptation strategy (NA dose) was compared with a
 231 fully manual approach (MA dose) based on the clinical goals defined in the
 232 PROTECT protocol. Robustness evaluations for the NA dose were computed
 233 for XT and PT using different residual setup errors of 2, 4, and 5 mm on the
 234 iCTV [28, 30–32], and were compared to the 7 mm setup error used for the
 235 MA dose. For PT, both NA and MA doses evaluation also included a 2.6%
 236 range uncertainty. We assessed statistically significant differences ($p = 0.05$)
 237 between the MA and NA doses in both nominal and worst-case scenarios.
 238 This analysis highlights which metrics are most affected when adapting EC
 239 patients and illustrates the potential benefits of adaptation.

Results

Adaptation rate

Figure 3 shows the PROTECT criteria for adaptation after three weeks of treatment in the nominal and worst-case scenarios. It quantifies the adaptation need for each patient, across both modalities and different setup margins. In Figure 3a for XT, ten patients (55.6 %) would require adaptation with a 5 mm setup margin, of which four (22.2 %) due to an insufficient target coverage in any scenario and only one (5.6 %) due to an insufficient target coverage in the nominal dose. The remaining six patients required adaptation due to an excess in body $D7\text{cm}^3$. With 4 mm and 2 mm setup margins, eight patients (44.4 %) would require adaptation, of which only one was due to insufficient target coverage. For PT, three patients (16.7 %) would require adaptation with 5 mm, 4 mm, and 2 mm setup margins, all due to insufficient target coverage. Only two patients would require adaptation if only the nominal dose is considered. All these results are summarized in Table 3.

In Table 3, we observed a clear impact of the triggering criteria on the adaptation rate despite being very similar in terms of dosimetric requirements. Interestingly, the adaptation rate is highly dependent on the selection of the robustness evaluation method, which may involve either the voxelwise minimum dose or the worst-case dose [13]. We also observed that the choice of residual setup error has minimal impact on the adaptation rate estimated by computing the iCTV V95% on the worst-case dose (PROTECT), except a higher rate for a 5mm margin in IMRT. This is likely due to the severe underdosage observed on the NA worst-case scenario dose for patients 3 and 18, as shown in Figure 3 (e.g., iCTV V95% $\sim$ 90% or lower for IMPT with a 5 mm residual setup). However, when estimating the adaptation rate using iCTV D98% on the voxelwise minimum dose [15, 22, 23], the adaptation rate drops from 72.2% to 33.3% for IMRT and from 72.2% to 44.4% for IMPT, when decreasing the margin from 5 mm to 2 mm.

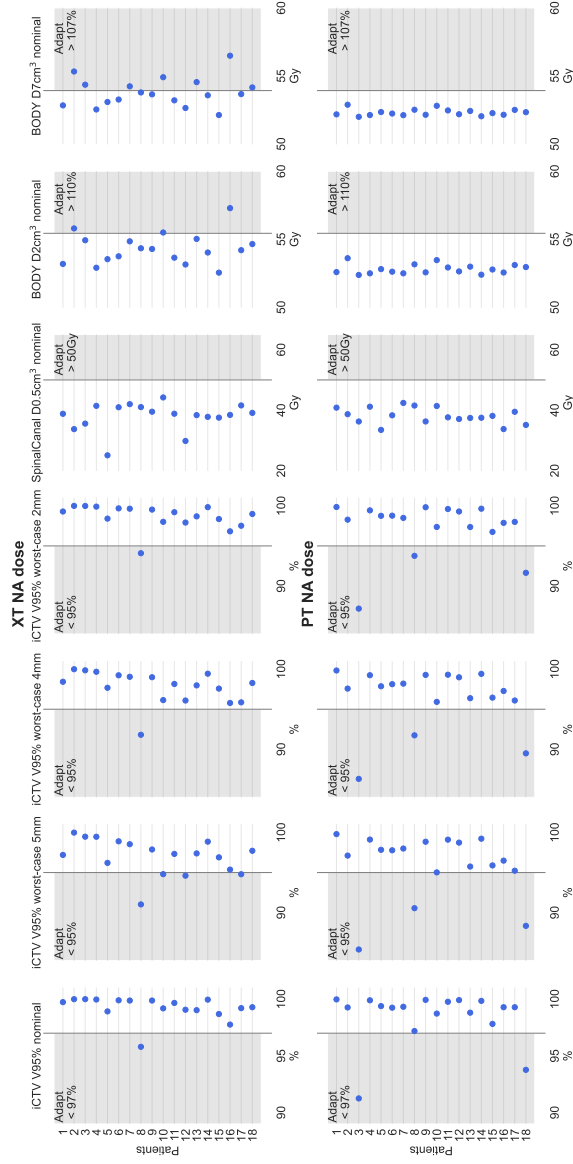


Figure 3: Criteria for adaptation in the nominal case for iCTV and organs at risk (OARs) and in the worst-case scenario computed on the iCTV with a 2, 4 or 5 mm residual setup errors. The range uncertainty for proton plans evaluation is 2.6%. DVH metrics are computed over the non-adapted dose (NA dose). Thresholds for each DVH metrics are indicated in the upper corners of each subfigure, % thresholds for dose metrics refer to % of prescription (50.4 Gy).

Table 3: Adaptation rate resulting from PROTECT’s and Table 1 criteria, applied on the NA dose. Parentheses indicate the structure on which the criteria are evaluated. iCTV and iPTV correspond to the aggregation of CTV and PTV over all 4D-CT phases, respectively. Worst-case stands for the worst-case scenario.

Modality	Criteria	Adaptation rate (%)
IMRT	PROTECT 15 th fraction (iCTV + OARs)	
	<i>Nominal</i>	8/18 (44.4%)
	<i>Worst-case 5mm</i>	10/18 (55.6%)
	<i>Worst-case 4mm</i>	8/18 (44.4%)
	<i>Worst-case 2mm</i>	8/18 (44.4%)
	PROTECT 15 th fraction (iCTV)	
	<i>Nominal</i>	1/18 (5.6%)
	<i>Worst-case 5mm</i>	4/18 (22.2%)
	<i>Worst-case 4mm</i>	1/18 (5.6%)
	<i>Worst-case 2mm</i>	1/18 (5.6%)
	Nyeng et al. [20] (iCTV + iPTV)	
	<i>Nominal</i>	11/18 (61.1%)
	Visser et al. [22] & Anakotta et al. [15] & Canter el al. [23] (iCTV)	
	<i>Voxelwise minimum 5mm</i>	13/18 (72.2%)
	<i>Voxelwise minimum 4mm</i>	11/18 (61.1%)
	<i>Voxelwise minimum 2mm</i>	6/18 (33.3%)
IMPT	PROTECT 15 th fraction (iCTV + OARs)	
	<i>Nominal</i>	2/18 (11.1%)
	<i>Worst-case 5mm</i>	3/18 (16.7%)
	<i>Worst-case 4mm</i>	3/18 (16.7%)
	<i>Worst-case 2mm</i>	3/18 (16.7%)
	PROTECT 15 th fraction (iCTV)	
	<i>Nominal</i>	2/18 (11.1%)
	<i>Worst-case 5mm</i>	3/18 (16.7%)
	<i>Worst-case 4mm</i>	3/18 (16.7%)
	<i>Worst-case 2mm</i>	3/18 (16.7%)
	Nyeng et al. [20] (iCTV)	
	<i>Nominal</i>	6/18 (33.3%)
	Visser et al. [22] & Anakotta et al. [15] & Canter el al. [23] (iCTV)	
	<i>Voxelwise minimum 5mm</i>	13/18 (72.2%)
	<i>Voxelwise minimum 4mm</i>	12/18 (66.7%)
	<i>Voxelwise minimum 2mm</i>	8/18 (44.4%)

269 **Benefits of adaptation for EC**

270 For each patient requiring adaptation according to PROTECT's criteria, a
271 new adapted plan was created. This includes 8 to 10 patients in IMRT and
272 2 to 3 patients in IMPT, depending on the residual setup error considered
273 for evaluation (see 1). This experiment highlights how these patients would
274 benefit from adaptation. Manual plan adaptation restored an acceptable tar-
275 get coverage (iCTV V95% > 99% in the nominal scenario and > 97% in the
276 worst-case scenario, see Supplementary Materials, Table S2) in IMRT and
277 IMPT. Figure 4 illustrates the general trend over this patient cohort. Statisti-
278 cally significant differences between the MA dose and the NA dose are also
279 highlighted in Figure 4. We clearly observe that when adaptation is required,
280 the MA dose achieves a higher median target coverage, and that the benefit
281 of adaptation is, on average, reduced when the NA dose is evaluated with a
282 lower residual setup error. This trend is even more pronounced for IMPT. In
283 the nominal scenario, the MA dose provides higher target coverage than the
284 NA dose in all patients, with differences in iCTV V95% ranging from [0% -
285 4.2%] for IMRT and from [2.8% - 8.8%] for IMPT. Moreover, the nominal dose
286 to most OARs is, on average, higher in the NA dose than in the MA dose
287 for both modalities. However, in the worst-case IMRT scenarios, we observe
288 that the MA dose yields a lower target coverage than the NA dose for two
289 patients (with differences up to 1.4%). This occurs for the iCTV V95% when
290 evaluated with 5, 4, and 2 mm residual setup errors. For the OARs, there
291 is a general trend where decreasing the residual setup error results in a NA
292 dose that reaches a lower dose to OARs in the worst-case scenarios. In some
293 cases, the NA dose to OARs becomes even lower than the MA dose.

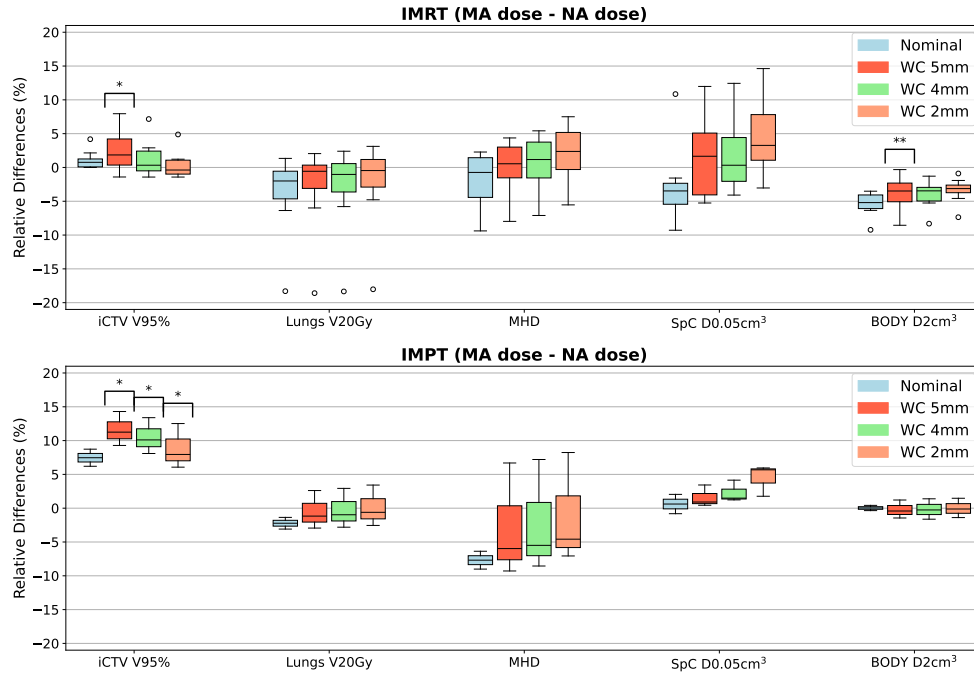


Figure 4: Comparison between MA dose and NA dose for patients requiring adaptation according to PROTECT's criteria (8 for IMRT nominal, 4 mm and 2 mm, 10 for IMRT 5 mm, 2 for IMPT nominal, 3 for IMPT 5 mm, 4 mm, and 2 mm). Boxplots show the difference (MA dose – NA dose) relative to the prescription dose (50.4 Gy) or to the structure's volume. Worst-case scenarios are computed with residual setup errors of 2 mm, 4 mm or 5 mm for the NA dose and a setup error of 7 mm for the MA dose. * indicates a statistically significant difference ($p = 0.05$) between MA dose and NA dose, ** indicates a highly statistically significant difference ($p = 0.001$) between MA dose and NA dose.

Discussion and conclusion

Adapting a plan during treatment requires significant computational and time resources. Therefore, assessing the need for adaptation is crucial to ensure patient comfort, treatment efficiency, and turnover within the treatment center. In order to assist clinicians with an automated decision tool, protocol for adaptation should rely on specific dosimetric criteria. If these clinical goals are not met with the dose recomputed on the anatomy of the

301 day, the treatment should be adapted or at least the dose distribution should
302 be visually inspected by a radiation oncologist and a decision should be made
303 together with a physicist.

304 Several adaptation criteria have been reported in the literature, often
305 based on local clinical experience rather than standardized guidelines or
306 consensus. Consequently, implementation of adaptive radiotherapy varies
307 across institutions. Our results show that the choice of criteria significantly
308 affects adaptation rates, in our EC patient cohort. These findings highlight
309 the need for standardized methodologies to ensure more consistent clinical
310 practice. The aim of our work is not to propose specific criteria, but to
311 illustrate how different criteria and residual setup errors influence adaptation
312 rates and to emphasize the importance of moving toward a consensus
313 in defining these criteria. It is noteworthy that although our database of
314 repeated CTs is not biased (all 18 patients underwent a repeated CT scan
315 at week 3 without any a priori selection), we limit our analysis to week 3.
316 Therefore, the adaptation rate that we report in this study might be different
317 if estimated in repeated CTs acquired during other weeks. The reported
318 adaptation rate may also suffer from other influences that are hard to quantify.
319 For instance, Table 1 cites a study that reporting that soft-tissue matching
320 is associated with a reduction of the adaptation rate. Our methodology is
321 also based on soft-tissue matching. As a consequence, our adaptation rates
322 may not be representative of what would be obtained in other clinics that use
323 bone-tissue matching. However, the focus of our present study is more about
324 the sensitivity of the adaptation rate to various parameters (the definition of
325 the adaptation rate, the residual error employed...) rather than the actual
326 value of the adaptation rate. This value will depend on various methodological
327 choices like the time of acquisition of the 4DCT, the explicit dosimetry
328 evaluation of breathing motion, soft-tissue or bone-tissue matching, etc.

329 The investigated adaptation thresholds come from published studies and
330 are mainly based on the evaluation of the target coverage. Despite representing
331 a very thin window in terms of target coverage requirements, they still
332 yielded a wide range of adaptation rates (e.g., for IMPT from 72.2% following
333 the protocol of Visser et al. [22] to 16.7% following PROTECT guidelines,
334 using a setup margin of 5 mm in both cases). We should also emphasize
335 the impact of the chosen robustness evaluation method as a critical factor
336 influencing the adaptation rate, as the chosen evaluation dose might be correlated
337 differently with target coverage [13]. In the voxelwise worst-dose

338 evaluation, different voxels may receive their worst dose in different sce-
339 narios. This theoretical dose is therefore unrealizable and potentially overly
340 conservative [29]. In contrast, the worst-case dose considers only the worst
341 among realizable scenarios, accounting for setup and possibly density shifts.

342 The observed discrepancies in adaptation rates between nearly equiva-
343 lent dosimetric criteria emphasize the need for consensus guidelines to stan-
344 dardize plan adaptation triggers in EC patients. Such guidelines would im-
345 prove consistency in clinical decision-making and ensure equitable treat-
346 ment quality across different institutions. However, achieving this consen-
347 sus is challenging due to the variability in patient anatomy, tumor location,
348 and clinical workflows. For example, tumors located at the gastroesophageal
349 junction are more prone to inter- and intra-fractional motion due to their
350 proximity to the diaphragm and lungs, necessitating frequent adaptations
351 when beams cross highly mobile regions such as the diaphragm or bowel
352 cavity [15, 20, 22]. In contrast, cervical esophageal cancer is more affected by
353 severe weight loss, leading to changes in patient geometry which can impact
354 differently the target coverage and surrounding OARs doses [33]. Establish-
355 ing robust guidelines would require data from large, diverse patient cohorts
356 that account for these anatomical and motion-related differences. Further-
357 more, variations in imaging and contouring protocols, and more generally in
358 radiotherapy workflows, as well as limited resources in many clinics, make
359 the implementation of consensus guidelines particularly challenging at this
360 stage. More extensive research and clinical experience are essential to bridge
361 these gaps and move closer to a standardized approach in adaptive radiother-
362 apy for EC patients.

363 In addition to diaphragm baseline shifts, bowel filling, gas, changes in
364 lung density, and esophagus displacement have also been reported as po-
365 tential causes of adaptation [10, 15, 20, 23]. Our patient cohort primarily
366 includes distal esophagus tumors (see Supplementary Materials, Table S4),
367 which are more likely to be affected by these anatomical changes [14]. In our
368 IMRT plans, we used two lateral-anterior beams (see Supplementary Mate-
369 rials, Table S1), with dose contributions crossing the diaphragmatic region.
370 These lateral dose contributions are generally reinforced due to the aware-
371 ness about heart sparing. Diaphragm position monitoring and informed
372 target alignment have been proposed to mitigate the impact of diaphragm
373 displacements [34]. The influence of bowel filling and gas may be miti-
374 gated by using only posterior beams, which also increase robustness against

375 respiratory motion [35]. Our IMPT treatment plans use only two poste-
376 rior beams (see Supplementary Materials, Table S1), partially explaining the
377 higher number of adaptation cases observed in IMRT compared to IMPT (e.g.,
378 61.1% for IMRT vs 33.2% for IMPT [20]).

379 The disparities in adaptation frequency, resulting from differing criteria,
380 raise the question of which guidelines are the most appropriate for determin-
381 ing when adaptation is required. This study primarily focuses on adaptation
382 criteria evaluating underdosage, but significant increases in doses to healthy
383 tissues might also trigger adaptation. Geometric criteria and visual inspec-
384 tion are also widely used, although they require human intervention and
385 cannot be automated. Besides, these geometric criteria do not always result
386 in a meaningful dosimetric impact after adaptation, as found by Möller et
387 al. [10]. Ideally, the dosimetric benefit of treatment plan adaptation should
388 translate into an improvement in tumor control probability (TCP), a decrease
389 in normal tissue complications [36–38] and improved quality of life. How-
390 ever, this approach relies on the existence of externally validated models.

391 The choice of the residual setup error for evaluation must be carefully
392 considered, as it directly influences the number of patients requiring adap-
393 tation. The reasoning behind the various options for residual setup uncer-
394 tainties lies in the fact that intra-fractional motion (other than breathing)
395 is not entirely captured in the newly acquired image. Additionally, unpre-
396 dicted isocenter shifts and delineation errors must be accounted for. **Those**
397 **uncertainties that occur after patient alignment entail systematic and ran-**
398 **dom errors, whose effects can be approximated by a global residual setup**
399 **margin (as it is the case during treatment plan optimization).** Our results
400 showed that PT was generally less sensitive to the choice of the setup eval-
401 uation margin, suggesting that our robustly optimized PT plans for EC pa-
402 tients may achieve treatment plans more robust to anatomical changes than
403 XT with PTV margins. However, these findings must be interpreted cau-
404 tiously. When evaluating iCTV V95% with PROTECT criteria, 3 out of 4
405 patients requiring adaptation for a 5 mm residual setup error in IMRT are
406 borderline and might not be adapted after visual inspection. In contrast, PT
407 patients requiring adaptation showed more significant target misses (V95%
408 < 90% for 2 patients with 5 and 4 mm residual setup errors). Furthermore,
409 in IMRT, the generally higher dose bath accounts for most adaptation cases,
410 while acceptable target coverage was still maintained.

411 Furthermore, the experiment comparing the MA dose with the NA dose
412 demonstrated that in the nominal scenario, adaptation may lead to a safer
413 target coverage and a lower dose of OAR for EC patients. However, only
414 a few of these improvements reach statistical significance, and results for
415 worst-case scenarios are sometimes worse for the MA dose. This might be
416 caused by the fixed margins used to optimize and evaluate adapted doses
417 in this work. Existing research suggests that ART with reduced margins
418 may enhance organ sparing without compromising target coverage [39, 40],
419 leading to a potential reduction in normal tissues complication probabilities
420 (NTCP) [41]. These findings indicate that, despite evidence for dose distri-
421 bution improvements, ART with fixed margins may not provide significant
422 clinical benefits, and clinical resources might be put to better use.

423 Typically, the benefits of adaptation are evaluated based on the dose accu-
424 mulated over rCTs. This provides relevant insight into the effect and neces-
425 sity of adaptation over the entire treatment course. A limitation of our study
426 is that our patients only had repeated scans after three weeks of treatment.
427 Therefore, the benefits of adaptation were evaluated based on the adapted
428 non-accumulated dose to avoid the expected effects of replanning being lost
429 in accumulation. We estimated that in most adaptive strategies, patients will
430 undergo more than one weekly rCT, and the benefits of adaptation may be
431 better observed on the accumulated dose in such cases. Although the use of
432 an rCT at week 3 is not ideal, as the effects of anatomical changes should
433 be mitigated as soon as possible [22], several studies still report triggered
434 adaptation based on late-acquired images [23, 42].

435 Even though some studies also recommend the robust evaluation of the
436 NA dose with respect to breathing motion [20, 21], our evaluations were
437 only conducted on the average rCT. We assumed that robustness evaluation
438 on the average rCT was sufficient to detect target underdosage, in line with
439 the findings of Visser et al. [43] and Canters et al. [23]. **However, despite**
440 **those publications, it is clear that a thorough 4D evaluation of the treatment**
441 **plans, including the interplay effect, would offer the best guarantees.**

Conclusion

In conclusion, this work provides insights into the dependence of adaptation rates for EC patients on both the adaptation criteria and the extent of residual setup error considered during evaluation. Our results highlight that the choice of adaptation criteria can significantly influence which patients require plan adaptation, reinforcing the need for standardized methodologies or at least a consensus framework. Consensus protocols should also consider tumor location on the esophageal wall, as typical anatomical changes may affect these regions differently. **Considering all relevant dose metrics, the adaptation rate was higher for IMRT than IMPT, because of the sensitivity of the body D7cm³ metric. However, if we consider target coverage only, IMPT features a slightly higher sensitivity compared to IMRT.** However, the number of required adaptations can potentially be mitigated by selecting an appropriate beam configuration and residual setup error. Additionally, the comparatively small benefits obtained from adapting the treatment plan for specific patients further motivate the implementation of adaptive strategies with reduced margins.

459 **Supplementary materials**

460 **PROTECT planning protocol**

461 **IMRT:** Plans were made using 7 beams on the average CT: one purely
462 anterior, 2 posterior obliques and 4 anterior obliques (beam angles: 0°, 155°,
463 220°, 30°, 330°, 60°, 300°). The prescribed dose of 50.4 Gy was divided into 28
464 fractions of 1.8 Gy each. To cover for possible setup errors and variability in
465 breathing motion, a 7mm iPTV was also created and the dose was optimized
466 on this iPTV target. Robust evaluation was performed on Average CT, re-
467 calculating the dose for a 7 mm setup error and a 0% range error. Fourteen
468 scenarios were evaluated in total. The nominal dose was recomputed on all
469 breathing phases as well (between 8 and 10 images, depending on the pa-
470 tient). The treatment machine is Varian Halcyon, which delivers the dose
471 using a 6MV FFF (flattening filter free) photon beam. All photon plans were
472 optimized in RayStation 2023B and computed with a Monte Carlo v3.0 dose
473 engine, both at optimization time and for final dose calculation.

474 **IMPT:** Plans were made using 2 oblique posterior beams on the average
475 CT (beam angles: 150° and 180°). The prescribed dose of 50.4 Gy was divided
476 into 28 fractions of 1.8 Gy each. Robust worst-case optimization [29] was
477 performed in the iCTV target, using 7 mm setup error and 2.6% range error.
478 Robust evaluation was performed on Average CT (28 scenarios in total, due
479 to non-zero range error) and all the breathing phases, as well. The machine
480 used for treatment is an IBA Proteus Plus. All proton plans were optimized
481 in RayStation 2023B and computed using a Monte Carlo v5.3 dose engine,
482 both at optimization time and for final dose calculation.

Modality	Direction	Gantry Angle
IMRT (7 beams)	Anterior	0°
	Left Anterior Oblique	30°
	Left Anterior Oblique	60°
	Left Posterior Oblique	155°
	Right Posterior Oblique	220°
	Right Anterior Oblique	300°
	Right Anterior Oblique	330°
IMPT	Left Posterior Oblique	150°
	Right Posterior Oblique	180°

Table S1: Beam configuration used in our treatment planning protocols for IMRT and IMPT.

First priority constraints		
Target	Dose (nominal plan)	Dose (all scenarios)
iCTV	$V95\%_{iCTV} > 99\%$	$V95\%_{iCTV} > 97\%$
iCTV	$98\% < D_{mean} < 102\%$	
Second priority constraints		
Target	Dose (nominal plan)	Dose (all scenarios)
iCTV		$V95\%_{iCTV} > 98\%$
iCTV	$99\% < D_{mean} < 101\%$	

Table S2: PROTECT target dose constraints.

First priority constraints		
Organs at risk	Dose (nominal plan)	Dose (all scenarios)
Spinal cord	$D_{0.05cm^3} < 45Gy$	
Spinal cord PRV	$D_{0.05cm^3} < 50Gy$	$D_{0.05cm^3} < 50Gy$
Lungs	$MLD < 20Gy$ $V_{20Gy} < 35\%$ $V_{5Gy} < 70\%$	<i>Not evaluated</i>
Body	$D_{0.05cm^3} < 110\%$ $D_{1cm^3} < 107\%$ $D_{5cm^3} < 107\%$	$D_{1cm^3} < 110\%$
Second priority constraints		
Organs at risk	Dose (nominal plan)	Dose (all scenarios)
Lungs	$V_{5Gy} < 60Gy$	<i>Not evaluated</i>
Heart	$V_{40Gy} < 30\%$ $V_{25Gy} < 50\%$ $MHD < 26Gy$	<i>Not evaluated</i>

Table S3: PROTECT dose constraints for OARs.

		Number	Percentage
Gender	Male	18	100%
	Female	0	0%
Age (years)	Median (range)	64 [40 – 77]	
BMI	Median (range)	27 [23 – 38]	
Histology	ACC	13	72.2%
	SCC	4	22.2%
	Other	1	5.6%
Tumor location	Mid-thoracic	3	16.7%
	Low-thoracic	12	66.7%
	GEJ	3	16.7%
TN stage	T1N1	1	5.6%
	T2N1	1	5.6%
	T3N0	1	5.6%
	T3N1	5	27.8%
	T3N2	6	33.3%
	T3N3	1	5.6%
	Unknown	3	16.7%
GTV volume (cm³)	Median (range)	47.5 [11.2 – 138.9]	

Table S4: Patient characteristics.

	Fraction 1-10	Fraction 11-21	Fraction 21-end#
iCTV			
<i>Nominal plan</i>	$V95\%_{iCTV} >98\%$	$V95\%_{iCTV} >97\%$	$V95\%_{iCTV} >96\%$
<i>All scenarios</i>	$V95\%_{iCTV} >96\%$	$V95\%_{iCTV} >95\%$	$V95\%_{iCTV} >94\%$
Spinal cord			
<i>Nominal plan</i>	$D_{0.05cm^3} <50Gy$	$D_{0.5cm^3} <50Gy$	$D_{1cm^3} <50Gy$
Body			
<i>Nominal plan</i>	$D_{1cm^3} <110\%$	$D_{2cm^3} <110\%$	$D_{3cm^3} <110\%$
	$D_{5cm^3} <107\%$	$D_{7cm^3} <107\%$	$D_{9cm^3} <107\%$

Table S5: Dose Constraints on iCTV and OARs, evaluated on the treatment plan recalculated on the new average 4D-CT. If these tolerances are exceeded, an adaptive plan should be made. The effect of residual setup error and range uncertainty on the iCTV should be evaluated using 2 mm setup error and the same range uncertainty as used for the initial treatment plan (2.6% in our case.)

	2 mm	4 mm	5 mm	7 mm
Photons				
$\rho = 0\%$	(-2.0,0,0)	(-4.0,0,0)	(-5.0,0,0)	(-7.0,0,0)
	(2.0,0,0)	(4.0,0,0)	(5.0,0,0)	(7.0,0,0)
	(0,2.0,0)	(0,4.0,0)	(0,5.0,0)	(0,7.0,0)
	(0,-2.0,0)	(0,-4.0,0)	(0,-5.0,0)	(0,-7.0,0)
	(0,0,-2.0)	(0,0,-4.0)	(0,0,-5.0)	(0,0,-7.0)
	(0,0,2.0)	(0,0,4.0)	(0,0,5.0)	(0,0,7.0)
	(-1.2,1.2,-1.2)	(-2.3,2.3,-2.3)	(-2.9,2.9,-2.9)	(-4.0,4.0,-4.0)
	(-1.2,-1.2,-1.2)	(-2.3,-2.3,-2.3)	(-2.9,-2.9,-2.9)	(-4.0,-4.0,-4.0)
	(1.2,1.2,-1.2)	(2.3,2.3,-2.3)	(2.9,2.9,-2.9)	(4.0,4.0,-4.0)
	(1.2,-1.2,-1.2)	(2.3,-2.3,-2.3)	(2.9,-2.9,-2.9)	(4.0,-4.0,-4.0)
	(1.2,-1.2,1.2)	(2.3,-2.3,2.3)	(2.9,-2.9,2.9)	(4.0,-4.0,4.0)
	(-1.2,-1.2,1.2)	(-2.3,-2.3,2.3)	(-2.9,-2.9,2.9)	(-4.0,-4.0,4.0)
	(-1.2,1.2,1.2)	(-2.3,2.3,2.3)	(-2.9,2.9,2.9)	(-4.0,4.0,4.0)
	(1.2,1.2,1.2)	(2.3,2.3,2.3)	(2.9,2.9,2.9)	(4.0,4.0,4.0)
Protons				
$\rho = \pm 2.6\%$	± 2.6 (-2.0,0,0)	± 2.6 (-4.0,0,0)	± 2.6 (-5.0,0,0)	± 2.6 (-7.0,0,0)
	± 2.6 (2.0,0,0)	± 2.6 (4.0,0,0)	± 2.6 (5.0,0,0)	± 2.6 (7.0,0,0)
	± 2.6 (0,2.0,0)	± 2.6 (0,4.0,0)	± 2.6 (0,5.0,0)	± 2.6 (0,7.0,0)
	± 2.6 (0,-2.0,0)	± 2.6 (0,-4.0,0)	± 2.6 (0,-5.0,0)	± 2.6 (0,-7.0,0)
	± 2.6 (0,0,-2.0)	± 2.6 (0,0,-4.0)	± 2.6 (0,0,-5.0)	± 2.6 (0,0,-7.0)
	± 2.6 (0,0,2.0)	± 2.6 (0,0,4.0)	± 2.6 (0,0,5.0)	± 2.6 (0,0,7.0)
	± 2.6 (-1.2,1.2,-1.2)	± 2.6 (-2.3,2.3,-2.3)	± 2.6 (-2.9,2.9,-2.9)	± 2.6 (-4.0,4.0,-4.0)
	± 2.6 (-1.2,-1.2,-1.2)	± 2.6 (-2.3,-2.3,-2.3)	± 2.6 (-2.9,-2.9,-2.9)	± 2.6 (-4.0,-4.0,-4.0)
	± 2.6 (1.2,1.2,-1.2)	± 2.6 (2.3,2.3,-2.3)	± 2.6 (2.9,2.9,-2.9)	± 2.6 (4.0,4.0,-4.0)
	± 2.6 (1.2,-1.2,-1.2)	± 2.6 (2.3,-2.3,-2.3)	± 2.6 (2.9,-2.9,-2.9)	± 2.6 (4.0,-4.0,-4.0)
	± 2.6 (1.2,-1.2,1.2)	± 2.6 (2.3,-2.3,2.3)	± 2.6 (2.9,-2.9,2.9)	± 2.6 (4.0,-4.0,4.0)
	± 2.6 (-1.2,-1.2,1.2)	± 2.6 (-2.3,-2.3,2.3)	± 2.6 (-2.9,-2.9,2.9)	± 2.6 (-4.0,-4.0,4.0)
	± 2.6 (-1.2,1.2,1.2)	± 2.6 (-2.3,2.3,2.3)	± 2.6 (-2.9,2.9,2.9)	± 2.6 (-4.0,4.0,4.0)
	± 2.6 (1.2,1.2,1.2)	± 2.6 (2.3,2.3,2.3)	± 2.6 (2.9,2.9,2.9)	± 2.6 (4.0,4.0,4.0)

Table S6: Scenario groups used for robustness evaluation. Columns correspond to setup uncertainty magnitudes (2, 4, 5, and 7 mm). Photon scenarios include isotropic setup shifts only ($\rho = 0\%$), while proton scenarios combine setup shifts with density scaling ($\rho = \pm 2.6\%$). Shifts are expressed in millimetres in the patient coordinate system: Right–Left (R–L), Inferior–Superior (I–S), and Anterior–Posterior (A–P).

References

- [1] Steven H. Lin, Ritsuko Komaki, Zhongxing Liao, Caimiao Wei, Bevan Myles, Xiaomao Guo, Matthew Palmer, Radhe Mohan, Stephen G. Swisher, Wayne L. Hofstetter, Jaffer A. Ajani, and James D. Cox. Proton Beam Therapy and Concurrent Chemotherapy for Esophageal Cancer. *International Journal of Radiation Oncology*Biology*Physics*, 83(3):e345–e351, July 2012.
- [2] Jingya Wang, Caimiao Wei, Susan L. Tucker, Bevan Myles, Matthew Palmer, Wayne L. Hofstetter, Stephen G. Swisher, Jaffer A. Ajani, James D. Cox, Ritsuko Komaki, Zhongxing Liao, and Steven H. Lin. Predictors of Postoperative Complications After Trimodality Therapy for Esophageal Cancer. *International Journal of Radiation Oncology*Biology*Physics*, 86(5):885–891, August 2013.
- [3] Steven H. Lin, Kenneth W. Merrell, Jincheng Shen, Vivek Verma, Arlene M. Correa, Lu Wang, Peter F. Thall, Neha Bhooshan, Sarah E. James, Michael G. Haddock, Mohan Suntharalingam, Minesh P. Mehta, Zhongxing Liao, James D. Cox, Ritsuko Komaki, Reza J. Mehran, Michael D. Chuong, and Christopher L. Hallemeier. Multi-institutional analysis of radiation modality use and postoperative outcomes of neoadjuvant chemoradiation for esophageal cancer. *Radiotherapy and Oncology*, 123(3):376–381, June 2017.
- [4] H.R. Mortensen, P. Populaire, L. Hoffmann, D.S. Moeller, A. Appelt, P. Naftoux, C.T. Muijs, C. Grau, M.A. Hawkins, E.G.C. Troost, G. Defraene, R. Canters, C.S. Clarke, D.C. Weber, E.W. Korevaar, K. Haustermans, M. Nordsmark, Val Gebski, M.P. Achiam, Sheraz R. Markar,

- 509 Ganesh Radhakrishna, Maaïke Berbee, Daniele Scartoni, Ester Or-
510 landi, Jerome Doyen, Vincent Gregoire, Gilles Crehange, Johannes Lan-
511 gendijk, Paula Lorgelly, Hedwig M. Blommenstein, Camilla S. Byskov,
512 Mai L. Ehmsen, Maria Fuglsang Jensen, Gloria Vilches Freixas, and Re-
513 becca Bütof. Proton versus photon therapy for esophageal cancer –
514 A trimodality strategy (PROTECT) NCT050555648. *Radiotherapy and*
515 *Oncology*, 190:109980, January 2024.
- 516 [5] Jan-Jakob Sonke, Marianne Aznar, and Coen Rasch. Adaptive Ra-
517 diotherapy for Anatomical Changes. *Seminars in Radiation Oncology*,
518 29(3):245–257, July 2019.
- 519 [6] David L. Schwartz, Adam S. Garden, Jimmy Thomas, Yipei Chen, Yong-
520 bin Zhang, Jan Lewin, Mark S. Chambers, and Lei Dong. Adaptive Ra-
521 diotherapy for Head-and-Neck Cancer: Initial Clinical Outcomes From
522 a Prospective Trial. *International Journal of Radiation Oncology, Biology,*
523 *Physics*, 83(3):986–993, July 2012.
- 524 [7] Haihua Yang, Wei Hu, Wei Wang, Peifang Chen, Weijun Ding, and Wei
525 Luo. Replanning During Intensity Modulated Radiation Therapy Im-
526 proved Quality of Life in Patients With Nasopharyngeal Carcinoma.
527 *International Journal of Radiation Oncology*Biological*Physics*, 85(1):e47–
528 e54, January 2013.
- 529 [8] Ditte Sloth Møller, Christina Maria Lutz, Azza Ahmed Khalil, Markus
530 Alber, Marianne Ingerslev Holt, Maria Kandi, Hjørdis Hjalting Schmidt,
531 Marie Tvillum, Ane Appelt, Marianne Marquard Knap, and Lone Hoff-
532 mann. Survival benefits for non-small cell lung cancer patients treated
533 with adaptive radiotherapy. *Radiotherapy and Oncology*, 168:234–240,
534 March 2022.
- 535 [9] Margriet Kwint, Sanne Conijn, Eva Schaake, Joost Kneijens, Maddalena
536 Rossi, Peter Remeijer, Jan-Jakob Sonke, and Jose Belderbos. Intra tho-
537 racic anatomical changes in lung cancer patients during the course of
538 radiotherapy. *Radiotherapy and Oncology*, 113(3):392–397, December
539 2014.
- 540 [10] Ditte Sloth Møller, Marianne Ingerslev Holt, Markus Alber, Marie
541 Tvillum, Azza Ahmed Khalil, Marianne Marquard Knap, and Lone Hoff-
542 mann. Adaptive radiotherapy for advanced lung cancer ensures tar-

- 543 get coverage and decreases lung dose. *Radiotherapy and Oncology*,
544 121(1):32–38, October 2016.
- 545 [11] Marcel Van Herk, Peter Remeijer, Coen Rasch, and Joos V. Lebesque.
546 The probability of correct target dosage: Dose-population histograms
547 for deriving treatment margins in radiotherapy. *International Journal*
548 *of Radiation Oncology*Biophysics*, 47(4):1121–1135, July 2000.
- 549 [12] ICRU Report 62, Prescribing, Recording and Reporting Photon Beam
550 Therapy (Supplement to ICRU 50) – ICRU.
- 551 [13] Erik W. Korevaar, Steven J. M. Habraken, Daniel Scandurra, Roel G. J.
552 Kierkels, Mirko Unipan, Martijn G. C. Eenink, Roel J. H. M. Steen-
553 bakkens, Stephanie G. Peeters, Jaap D. Zindler, Mischa Hoogeman, and
554 Johannes A. Langendijk. Practical robustness evaluation in radiother-
555 apy – A photon and proton-proof alternative to PTV-based plan evalu-
556 ation. *Radiotherapy and Oncology*, 141:267–274, December 2019.
- 557 [14] Lone Hoffmann, Per R. Poulsen, Thomas Ravkilde, Jenny Bertholet,
558 Iryna Kruhlikava, Bjørn L. Helbo, Mai L. Schmidt, and Marianne Nords-
559 mark. Setup strategies and uncertainties in esophageal radiotherapy
560 based on detailed intra- and interfractional tumor motion mapping. *Radiotherapy and Oncology*, 136:161–168, July 2019.
- 562 [15] R. Melissa Anakotta, Hans P. van der Laan, Sabine Visser, Cassia O.
563 Ribeiro, Margriet Dieters, Johannes A. Langendijk, Stefan Both, Erik W.
564 Korevaar, Nanna M. Sijtsema, Antje Knopf, and Christina T. Muijs.
565 Weekly robustness evaluation of intensity-modulated proton therapy
566 for oesophageal cancer. *Radiotherapy and Oncology*, 151:66–72, Octo-
567 ber 2020.
- 568 [16] Melissa Thomas, Robin De Roover, Schalk van der Merwe, Maarten
569 Lambrecht, Gilles Defraene, and Karin Haustermans. The use of tumour
570 markers in oesophageal cancer to quantify setup errors and baseline
571 shifts during treatment. *Clinical and Translational Radiation Oncology*,
572 26:8–14, January 2021.
- 573 [17] E. Sterpin, L. Widesott, K. Poels, M. Hoogeman, E. W. Korevaar,
574 M. Lowe, S. Molinelli, and F. Fracchiolla. Robustness evaluation of pen-
575 cil beam scanning proton therapy treatment planning: A systematic
576 review. *Radiotherapy and Oncology*, 197:110365, August 2024.

- 577 [18] Harald Paganetti, Pablo Botas, Gregory C. Sharp, and Brian Winey.
578 Adaptive proton therapy. *Phys Med Biol*, 66(22), November 2021.
- 579 [19] Nienke Hoekstra, Steven Habraken, Annemarie Swaak-Kragten, Mis-
580 cha Hoogeman, and Jean-Philippe Pignol. Intrafraction motion during
581 partial breast irradiation depends on treatment time. *Radiotherapy and*
582 *Oncology*, 159:176–182, June 2021.
- 583 [20] Tine Bisballe Nyeng, Marianne Nordsmark, and Lone Hoffmann. Dosi-
584 metric evaluation of anatomical changes during treatment to identify
585 criteria for adaptive radiotherapy in oesophageal cancer patients. *Acta*
586 *Oncologica*, 54(9):1467–1473, October 2015.
- 587 [21] Muhammad Shamshad, Ditte Sloth Møller, Hanna Rahbek Mortensen,
588 Mai Lykkegaard Ehmsen, Maria Fuglsang Jensen, and Lone Hoffmann.
589 Bone versus soft-tissue setup in proton therapy for patients with oe-
590 sophageal cancer. *Acta Oncologica*, 61(8):994–1003, August 2022.
- 591 [22] Sabine Visser, Cássia O. Ribeiro, Margriet Dieters, Veronique E. Mul,
592 Anne G.H. Niezink, Arjen Van Der Schaaf, Antje-Christin Knopf, Jo-
593 hannes A. Langendijk, Erik W. Korevaar, Stefan Both, and Christina T.
594 Muijs. Robustness assessment of clinical adaptive proton and photon
595 radiotherapy for oesophageal cancer in the model-based approach. *Ra-*
596 *diotherapy and Oncology*, 177:197–204, December 2022.
- 597 [23] Richard Canters, Kim Van Der Klugt, Vicki Trier Taasti, Jeroen Buijsen,
598 Bastiaan Ta, Inge Steenbakkers, Ruud Houben, Gloria Vilches-Freixas,
599 and Maaïke Berbee. Robustness of intensity modulated proton treat-
600 ment of esophageal cancer for anatomical changes and breathing mo-
601 tion. *Radiotherapy and Oncology*, 198:110409, September 2024.
- 602 [24] Chaoqiong Ma, Zhen Tian, Ruoxi Wang, Zhongsu Feng, Fan Jiang,
603 Qiaoqiao Hu, Fang Yang, Anhui Shi, and Hao Wu. A prediction model
604 for dosimetric-based lung adaptive radiotherapy. *Medical Physics*,
605 49(10):6319–6333, 2022.
- 606 [25] Catarina Veiga, Guillaume Janssens, Ching-Ling Teng, Thomas Baudier,
607 Lucian Hotoiu, Jamie R. McClelland, Gary Royle, Liyong Lin, Lingshu
608 Yin, James Metz, Timothy D. Solberg, Zelig Tochner, Charles B. Simone,

- James McDonough, and Boon-Keng Kevin Teo. First Clinical Investigation of Cone Beam Computed Tomography and Deformable Registration for Adaptive Proton Therapy for Lung Cancer. *International Journal of Radiation Oncology*Biophysics*, 95(1):549–559, May 2016.
- [26] Karen Lim, James Stewart, Valerie Kelly, Jason Xie, Kristy K. Brock, Joanne Moseley, Young-Bin Cho, Anthony Fyles, Anna Lundin, Henrik Reh binder, Johan Löf, David A. Jaffray, and Michael Milosevic. Dosi metrically Triggered Adaptive Intensity Modulated Radiation Ther apy for Cervical Cancer. *International Journal of Radiation Oncol ogy*Biophysics*, 90(1):147–154, September 2014.
- [27] Sebastien A. A. Gros, Anand P. Santhanam, Alec M. Block, Bahman Emami, Brian H. Lee, and Cara Joyce. Retrospective Clinical Evaluation of a Decision-Support Software for Adaptive Radiotherapy of Head and Neck Cancer Patients. *Front. Oncol.*, 12, June 2022.
- [28] Melissa Thomas, Hanna R. Mortensen, Lone Hoffmann, Ditte S. Møller, Esther G. C. Troost, Christina T. Muijs, Maaïke Berbee, Rebecca Bütof, Owen Nicholas, Ganesh Radhakrishna, Gilles Defraene, Philippe Naf teux, Marianne Nordsmark, and Karin Haustermans. Proposal for the delineation of neoadjuvant target volumes in oesophageal cancer. *Ra diotherapy and Oncology*, 156:102–112, March 2021.
- [29] Albin Fredriksson, Anders Forsgren, and Björn Hårdemark. Minimax optimization for handling range and setup uncertainties in proton ther apy. *Medical Physics*, 38(3):1672–1684, 2011.
- [30] Hans Paul van der Laan, R. Melissa Anakotta, Erik W. Korevaar, Mar griet Dieters, J. Fred Ubbels, Robin Wijsman, Nanna M. Sijtsema, Stefan Both, Johannes A. Langendijk, Christina T. Muijs, and Antje C. Knopf. Organ sparing potential and inter-fraction robustness of adaptive in tensity modulated proton therapy for lung cancer. *Acta Oncologica*, 58(12):1775–1782, December 2019.
- [31] Cássia O. Ribeiro, Sabine Visser, Erik W. Korevaar, Nanna M. Sijtsema, R. Melissa Anakotta, Margriet Dieters, Stefan Both, Johannes A. Lan gendijk, Robin Wijsman, Christina T. Muijs, Arturs Meijers, and Antje Knopf. Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion: Robust

- 643 optimised plan evaluation by means of patient and machine specific
644 information. *Radiotherapy and Oncology*, 157:210–218, April 2021.
- 645 [32] Michelle Oud, Sebastiaan Breedveld, Marta Giżyńska, Michiel Kroesen,
646 Stefan Hutschemaekers, Steven Habraken, Steven Petit, Zoltán Perkó,
647 Ben Heijmen, and Mischa Hoogeman. An online adaptive plan library
648 approach for intensity modulated proton therapy for head and neck
649 cancer. *Radiotherapy and Oncology*, 176:68–75, November 2022.
- 650 [33] Sibel Karaca and Meltem Kırılı. Adaptive radiation therapy for cervical
651 esophageal cancer: Dosimetric and volumetric analysis. *J Gastrointest*
652 *Oncol*, 10(3):506–512, June 2019.
- 653 [34] Sabine Visser, Lydia A. den Otter, Cássia O. Ribeiro, Erik W. Korevaar,
654 Stefan Both, Johannes A. Langendijk, Christina T. Muijs, Nanna M. Si-
655 jtsema, and Antje Knopf. Diaphragm-Based Position Verification to
656 Improve Daily Target Dose Coverage in Proton and Photon Radiation
657 Therapy Treatment of Distal Esophageal Cancer. *International Journal*
658 *of Radiation Oncology*Biology*Physics*, 112(2):463–474, February 2022.
- 659 [35] Ditte Sloth Møller, Per Rugaard Poulsen, Andreas Hagner, Math-
660 ieu Dufour, Marianne Nordsmark, Tine Bisballe Nyeng, Hanna Rah-
661 bek Mortensen, Christina Maria Lutz, and Lone Hoffmann. Strate-
662 gies for Motion Robust Proton Therapy With Pencil Beam Scanning
663 for Esophageal Cancer. *International Journal of Radiation Oncol-*
664 *ogy*Biology*Physics*, 111(2):539–548, October 2021.
- 665 [36] Matthias Guckenberger, Anne Richter, Juergen Wilbert, Michael Flen-
666 tje, and Mike Partridge. Adaptive Radiotherapy for Locally Advanced
667 Non–Small-Cell Lung Cancer Does Not Underdose the Microscopic
668 Disease and has the Potential to Increase Tumor Control. *Interna-*
669 *tional Journal of Radiation Oncology*Biology*Physics*, 81(4):e275–e282,
670 November 2011.
- 671 [37] Molly M. McCulloch, Daniel G. Muenz, Matthew J. Schipper, Michael
672 Velec, Laura A. Dawson, and Kristy K. Brock. A simulation study to
673 assess the potential impact of developing normal tissue complication
674 probability models with accumulated dose. *Advances in Radiation On-*
675 *cology*, 3(4):662–672, October 2018.

- [38] Saba Ebrahimi and Gino J. Lim. A reinforcement learning approach for finding optimal policy of adaptive radiation therapy considering uncertain tumor biological response. *Artificial Intelligence in Medicine*, 121:102193, November 2021.
- [39] Simon van Kranen, Olga Hamming-Vrieze, Annelisa Wolf, Eugène Damen, Marcel van Herk, and Jan-Jakob Sonke. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. *International Journal of Radiation Oncology*Biophysics*, 96(3):653–660, November 2016.
- [40] Lena Nenoff, Michael Matter, Johanna Hedlund Lindmar, Damien Charles Weber, Antony John Lomax, and Francesca Albertini. Daily adaptive proton therapy – the key to innovative planning approaches for paranasal cancer treatments. *Acta Oncologica*, 58(10):1423–1428, October 2019.
- [41] Steven van de Water, Iris van Dam, Dennis R. Schaart, Abraham Al-Mamgani, Ben J. M. Heijmen, and Mischa S. Hoogeman. The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiotherapy and Oncology*, 120(1):56–62, July 2016.
- [42] Huixian Huang, Heming Lu, Guosheng Feng, Hailan Jiang, Jiaxin Chen, Jinjian Cheng, Qiang Pang, Zhiping Lu, Junzhao Gu, Luxing Peng, Shan Deng, Ying Mo, Danling Wu, and Yinglin Wei. Determining appropriate timing of adaptive radiation therapy for nasopharyngeal carcinoma during intensity-modulated radiation therapy. *Radiat Oncol*, 10(1):192, September 2015.
- [43] Sabine Visser, Erik W. Korevaar, Christina T. Muijs, Robin Wijsman, Johannes A. Langendijk, Pietro Pisciotta, Gabriel Gutierrez Marmitt, Cásia O. Ribeiro, and Stefan Both. Clinical 3D/4D cumulative proton dose assessment methods for thoracic tumours with large motion. *Radiotherapy and Oncology*, 182:109575, May 2023.