

Dosimetric comparison of distal esophageal carcinoma plans for patients treated with small-spot intensity-modulated proton versus volumetric-modulated arc therapies

Chenbin Liu¹ | Ronik S. Bhangoo¹ | Terence T. Sio¹ | Nathan Y. Yu¹ | Jie Shan¹ | Jennifer S. Chiang¹ | Julia X. Ding¹ | William G. Rule¹ | Shawn Korte¹ | Pedro Lara¹ | Xiaoning Ding¹ | Martin Bues¹ | Yanle Hu¹ | Todd DeWees² | Jonathan B. Ashman¹ | Wei Liu¹

¹Department of Radiation Oncology, Mayo Clinic, Phoenix, AZ, 85054, USA

²Division of Biostatistics, Mayo Clinic, Phoenix, AZ, 85054, USA

Author to whom correspondence should be addressed. Wei Liu
E-mail: liu.wei@mayo.edu.

Funding information

National Cancer Institute, Grant/Award Number: K25CA168984; Arizona Biomedical Research Commission; Lawrence W. and Marilyn W. Matteson Fund; Kemper Marley Foundation

Background: Esophageal carcinoma is the eighth most common cancer in the world. Volumetric-modulated arc therapy (VMAT) is widely used to treat distal esophageal carcinoma due to high conformality to the target and good sparing of organs at risk (OAR). It is not clear if small-spot intensity-modulated proton therapy (IMPT) demonstrates a dosimetric advantage over VMAT. In this study, we compared dosimetric performance of VMAT and small-spot IMPT for distal esophageal carcinoma in terms of plan quality, plan robustness, and interplay effects.

Methods: 35 distal esophageal carcinoma patients were retrospectively reviewed; 19 patients received small-spot IMPT and the remaining 16 of them received VMAT. Both plans were generated by delivering prescription doses to clinical target volumes (CTVs) on phase-averaged 4D-CT's. The dose-volume-histogram (DVH) band method was used to quantify plan robustness. Software was developed to evaluate interplay effects with randomized starting phases for each field per fraction. DVH indices were compared using Wilcoxon rank-sum test. For fair comparison, all the treatment plans were normalized to have the same $CTV_{high} D_{95\%}$ in the nominal scenario relative to the prescription dose.

Results: In the nominal scenario, small-spot IMPT delivered statistically significantly lower liver D_{mean} and $V_{30Gy[RBE]}$, lung D_{mean} , heart D_{mean} compared with VMAT. CTV_{high} dose homogeneity and protection of other OARs were comparable between the two treatments. In terms of plan robustness, the IMPT and VMAT plans were comparable for kidney $V_{18Gy[RBE]}$, liver $V_{30Gy[RBE]}$, stomach $V_{45Gy[RBE]}$, lung D_{mean} , $V_{5Gy[RBE]}$, and $V_{20Gy[RBE]}$, cord D_{max} and $D_{0.03cm^3}$, liver D_{mean} , heart $V_{20Gy[RBE]}$, and $V_{30Gy[RBE]}$, but IMPT was significantly worse for $CTV_{high} D_{95\%}$, D_{2cm^3} , and $D_{5\%}-D_{95\%}$, $CTV_{low} D_{95\%}$, heart D_{mean} , and $V_{40Gy[RBE]}$, requiring careful and experienced adjustments during the planning process and robustness considerations. The small-spot IMPT plans still met the standard clinical requirements after interplay effects were considered.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine.

Conclusions: Small-spot IMPT decreases doses to heart, liver, and total lung compared to VMAT as well as achieves clinically acceptable plan robustness. Our study supports the use of small-spot IMPT for the treatment of distal esophageal carcinoma.

KEY WORDS

distal esophageal, intensity-modulated proton therapy, interplay effects, small spot size, volumetric-modulated arc therapy

1 | INTRODUCTION

Esophageal carcinoma is the eighth most common cancer and the sixth most common cause of cancer deaths worldwide. There are an estimated 17 290 new cases and 15 850 deaths annually in America, and distal esophageal cancer cases are increasing rapidly in developed countries.^{1,2} In recent years, trimodality therapy (neoadjuvant chemoradiation followed by esophagectomy) has improved clinical outcomes in patients with locally advanced esophageal cancers compared to surgery alone.^{3,4} Concurrent chemotherapy, usually with weekly carboplatin and paclitaxel, combined with radiation doses of 41.4–50.4 Gy are considered standard treatments in the modern era.⁵ The long-term results of Radiation Therapy Oncology Group (RTOG) clinical trial 85-01 confirmed that chemoradiation increased overall survival for patients with esophageal carcinoma compared with radiotherapy (RT) alone.⁵ Due to proximity to surrounding organs at risk (OAR) such as heart, spinal cord, lungs, kidney, liver, and the remaining stomach, the RT planning for distal esophagus carcinoma poses special challenges.^{6,7} Sufficient radiation doses must be applied to the tumor and lymph node areas, while protecting nearby critical normal structures.

Volumetric-modulated arc therapy (VMAT) is an advanced form of intensity-modulated radiation therapy (IMRT) that can deliver a highly conformal dose distribution using single or multiple arcs.⁸ Compared with static-field IMRT, VMAT achieves similar OAR sparing and planning target volume (PTV) coverage with significantly shorter treatment time.^{9,10} Proton beam therapy delivers highly conformal target coverage, while sparing adjacent OARs due to its unique Bragg peak dose deposition characteristics. Proton beam therapy¹¹ has several forms of delivery including passive scattering (PSPT), uniform scanning, and intensity-modulated proton therapy (IMPT). Unlike PSPT, IMPT uses magnetic steering of a narrow proton beam, termed a *beamlet*, to deliver a modulated dose to a spot of a specified size, which offers improved high-dose conformality compared with PSPT and better OAR sparing in the mid- to low-dose range compared to IMRT.^{12,13} Therefore, it is hypothesized that IMPT can improve the therapeutic ratio which will result in fewer adverse effects, while achieving the same tumor control as IMRT or better.¹⁴

However, IMPT is highly sensitive to setup and range uncertainties, as well as vulnerable to respiratory motion present in the distal esophageal regions.^{15,16} The uncertainties originate from daily patient alignment, conversion of Hounsfield units to stopping power,

artifacts in computed tomography, and anatomical changes in patients etc.¹⁷ The interaction between dynamic beamlet delivery and respiratory motion, also called interplay effect, may degrade the quality of planned dose distributions,^{18–30} compromising the safety and efficacy of the proposed treatment. In addition, large spot sizes, common in older IMPT machines, tend to lead to larger penumbras, which results in undesired dose to adjacent OARs. The majority of new proton facilities offer smaller spot sizes, which can produce smaller penumbras and better OAR sparing. However, smaller spot-size plans can exacerbate the negative consequences of patient setup uncertainty and make interplay effects more prominent.^{25,31} As a result, we first need to quantify and then mitigate the impact of uncertainties and interplay effects for small-spot IMPT for the treatment of distal esophageal cancers.

Previous studies have focused on the comparison of plan quality alone among plans generated using either PSPT, large-spot IMPT (spot size σ as large as 6–15 mm depending on proton energies), and/or IMRT^{32–37} with no mention of plan uncertainties or motion effect analyses. Recently, Shiraishi et al. reported that in a large cohort of esophageal cancer patients, PSPT and large-spot IMPT significantly reduced radiation exposures to the whole heart and cardiac substructures compared with IMRT.³² Welsh et al. found that large-spot IMPT for distal esophageal carcinoma also lowered the doses to bilateral lungs and liver compared to IMRT.³⁷ More recently, a large, retrospective multi-institutional study also demonstrated that proton beam therapy appeared to be more clinically advantageous compared with 3-dimensional conformal RT and IMRT in lowering the incidence of pulmonary and cardiac complications as well as the mean length of in-hospital stay.³⁸

To the best of our knowledge, no dosimetric study has been reported comparing plan quality and plan robustness for small-spot IMPT and VMAT in the treatment of distal esophageal cancer. The aim of this study is to evaluate the feasibility of small spot IMPT for such treatments. We compared plan quality and robustness for VMAT and small-spot IMPT. The interplay effects of small-spot IMPT were also quantified.

2 | MATERIALS AND METHODS

2.A | Patient selection

Nineteen consecutive patients with distal esophageal carcinoma treated with IMPT and 16 patients treated with VMAT between

May 2014 and September 2017 at our institution were retrospectively reviewed. Table 1 shows the patient characteristics for IMPT and VMAT treatment groups. The patients included in this study were not randomly selected, but were carefully selected by an experienced physicist from the existing database to ensure that the patients from the two treatment groups did not show significant differences in age, gender, body mass, tumor volumes, motion amplitude, or prescription dose (Table 1); the cases were consecutively considered based on radiotherapy treatment start date. Patients were excluded from this study if they were less than 18 yr old and/or not treated with curative intent. All treatment plans included in this work were approved and delivered clinically. Patients were staged using PET/CTs. All patients had an Eastern Cooperative Oncology Group performance status of 2 or better. No patients in this study had implanted cardiac devices.

2.B | Patient simulation and contouring

For both IMPT and VMAT treatments, the processes for patient simulation and contouring were similar. Four-dimensional computed tomography (4D-CT) was used to simulate all patients in the supine position. At our clinic, we defined the respiratory motion amplitude by the largest displacement of tumor geometric center in the anterior–posterior (A-P), superior–inferior (S-I), and right–left (R-L) directions in all 4D breathing-phase CTs. Most commonly, the S-I direction demonstrated the largest motion amplitude. Immobilization devices including wing board (CIVCO, Kalona, IA) and vacuum bag (Elekta, Atlanta, GA) were used for all patients during 4D CT simulation and daily treatments.

TABLE 1 Patient characteristics between intensity-modulated proton therapy (IMPT) and volumetric-modulated arc therapy (VMAT) treatment groups.

	IMPT	VMAT	P-value
Patient number	19	16	
Age at treatment (yr)			0.79
Median (range)	71 (54–84)	71 (55–83)	
Gender			0.42*
Male, No. (%)	14 (73.7)	14 (87.5)	
Body mass (kg)			0.55
Median (range)	83.2 (53.8–123.0)	82.5 (69.9–111.3)	
CTV _{high} volume (cm ³)			0.40
Median (range)	205.0 (117.6–515.6)	219.0 (102.5–531.4)	
CTV _{low} volume (cm ³)			0.98
Median (range)	440.00 (151.8–1144.7)	447.10 (140.9–773.8)	
Motion amplitude (mm)			0.91
Median (range)	8 (4–13)	6.2 (6–13.7)	

CTV, clinical target volume.

*P-value from Fisher's exact test. Others are P-values from Wilcoxon rank-sum test.

The commercial treatment planning system (EclipseTM, version 13, Varian medical system, Palo Alto, CA) was used to generate treatment plans based on the simulation 4D CTs, which were used to localize the targets and OARs. Heart, cord, liver, stomach, bowel, and kidney and targets were contoured on the 4D-averaged CT. The same clinical target volumes (CTV_{high} and CTV_{low}) were used for IMPT and VMAT treatment planning. The prescription doses for CTV_{high} and CTV_{low} were 50.4 and 45.0 Gy[RBE] or 50.0 and 45.0 Gy[RBE] with simultaneously integrated boost, respectively; 41.4 Gy[RBE] was also allowed for CTV_{low} or if the patient was treated preoperatively. CTV_{low} and CTV_{high} were generated as follows: first, we identified the appropriate gross target volume (GTV) on the average CT or one of the respiratory phase CT scans, and the co-registered PET and CT scan. Then, a 3–4 cm expansion was added along the mucosal surface longitudinally, in addition to a 1–1.2 cm radial expansion for the CTVs which were anatomically constrained. The lower CTV volume typically included a small expansion of elective nodal volumes in the para-esophageal region. The treating radiation oncologist also adjusted the expansion of margins based on the pathology and location of the tumor; the potential microscopic tumor extent and anatomic boundaries of heart, lungs, liver, kidneys, and bowel were also taken into consideration in the final target delineation.

2.C | Treatment planning

In VMAT treatment planning, all treatment plans were generated on the averaged CT. We used planning target volumes (PTV), formed by a 5-mm uniform expansion of CTVs, for plan optimization and evaluation in VMAT. Most commonly, 2 to 3 arcs were used. Photon optimizer (PO) model in the EclipseTM TPS was used for VMAT optimization, and analytical anisotropic algorithm (AAA) model was used for dose calculation. The dose calculation grid size was 3 mm. The dosimetrists created treatment plans, which satisfied institutional dose constraints (Table 2) for OAR sparing. For target coverage, V_{100%} of PTV_{high} was at least 95% and D_{0.03cm³} of PTV_{high} was no more than 110% of prescription dose.

In IMPT treatment planning, all treatment plans were also generated on the averaged CT. An optimization target volume (OTV) was formed by uniform expansion of the CTV by 5 mm to help generate a robust plan. Proton spots were placed strategically outside of the OTV as well to ensure homogenous dose distributions in the OTV.

TABLE 2 Dose volume constraints for organs-at-risk for esophageal carcinoma treatments.

Structure	Dose limits (Gy[RBE])
Liver	D _{mean} < 25 Gy[RBE]; V _{30Gy[RBE]} < 60%
Total lung	V _{5Gy[RBE]} < 60%; V _{20Gy[RBE]} < 15%; D _{mean} < 15 Gy[RBE]
Spinal cord	D _{max} < 45 Gy[RBE]; V _{45Gy[RBE]} < 0.1%
Heart	V _{25Gy[RBE]} < 50%; V _{40Gy[RBE]} < 30%; D _{mean} < 26 Gy[RBE]
Left/right kidney	V _{18Gy[RBE]} < 10%; D _{mean} < 18 Gy[RBE]

Usually, 2 to 3 proton beams were used. Single field optimization (SFO) was always the first option, however, multiple-field optimization (MFO) was used if the plan generated by SFO did not meet the clinical requirements. For optimization, the pencil beam convolution supposition (PCS) model was used for SFO whereas the nonlinear uniform proton optimizer (NUPO) robust optimization was used for MFO in the Eclipse™ TPS to generate a robust MFO plan.^{39–43} PCS model was used for the final dose calculation and beam line modifiers for all IMPT plans. The dose calculation grid size was 2.5 mm. Dosimetrists and medical physicists chose beam angles to minimize the impact of motion and spare normal tissues. Density override (HU = 50) was used for the CTV in IMPT treatment planning to help generate plans robust to respiratory motion.^{30,44,45} After the plan was generated on the averaged CT, the dosimetrists created two verification dose distributions without the density override by recalculating the dose distributions on the maximum exhale and maximum inhale phases of the 4D CT. This allowed evaluation of the impact of respiratory motion. The original plan would be adjusted until the dose distributions calculated on the averaged CT, maximum exhale, and maximum inhale phases without density override all met required dose volume constraints (Table 2), plan robustness quantification thresholds, and the prescription criteria (see Robustness Quantification subsection). The dosimetrists were allowed to override the beam path around the diaphragm region to compensate for the change of water equivalent thickness due to respiratory motion. By adjusting the size of override region and its density values, the verification plans allowed for improvement while the original plan still met the clinical criterion.

The treating radiation oncologist approved the final treatment plan after careful evaluation of plan quality, plan robustness, and interplay effects. Plan robustness and interplay effect were only evaluated for IMPT planning (See details in Plan Quality Evaluation, Robustness Quantification, and Interplay Effects Evaluation subsections below). In addition, a second independent dose calculation was performed using an in-house developed Monte Carlo software.⁴⁶ We carefully commissioned our TPS by fudging the proton optics parameters,⁴⁷ so the dose distributions calculated by our TPS matched well with those generated by the Monte Carlo simulations.^{47,48} Based on 3 yr clinical experience at our institution, the deviation of CTV $D_{95\%}$ and D_{mean} between Monte Carlo simulation and TPS computation has been less than 3% and 2% for most of patients respectively.

2.D | Treatment delivery

The Clinac machines (Varian medical system, Palo Alto, CA) were used to deliver the VMAT plans. The related parameters including field information, energy, and estimated delivery duration are shown in Table S1. Typically, daily on-board imaging or cone-beam CT was used as the image-guided RT methods.

For IMPT, the Hitachi ProBeat-V spot-scanning proton beam machines (Hitachi, Tokyo, Japan) were used. The active scanning proton beam machine was commissioned to have an energy-dependent spot size (σ) of 2 to 6 mm, with a fixed spot spacing of 5 mm.

Our proton beam scanning machine had discrete proton energies ranging from 71.3 to 228.8 MeV. These discrete energies were carefully selected to minimize the ripples in the dose distribution along the beam direction and minimum MU effects. The energy layer switch time for all 97 energies ranged from 1.9 to 2.0 s, with an average of 1.91 s. The average spill length was 7.9 s. The average magnet preparation and verification time was 1.93 ms.⁴⁹ The related field and energy choices, estimated delivery duration, and repainting numbers are shown in the Table S2. We used orthogonal pair kV images to align to bony anatomy during the IMPT treatment.

For both IMPT and VMAT, the set up images were reviewed offline and approved by the treating radiation oncologist to make sure that the patient setup errors were within clinical tolerance.

2.E | Plan quality evaluation

Due to the different treatment planning methods used for VMAT and IMPT, the same target volumes, CTV_{high} and CTV_{low}, were used in plan quality, robustness, and interplay effects evaluation for fair comparison. CTV_{high} $D_{95\%}$ (the minimum dose covering the lowest 95% of the irradiated structure's volume), $D_{5\%}$, $D_{2\text{cm}^3}$ (the minimum dose to the highest irradiated 2cm³ of the structure's volume), and CTV_{low} $D_{95\%}$ were calculated from the CTV_{high} and CTV_{low} DVHs respectively. CTV $D_{95\%}$, $D_{5\%}$ - $D_{95\%}$, and $D_{2\text{cm}^3}$ were used to illustrate target dose coverage, target dose homogeneity, and hot spots respectively. The CTV-related DVH band width was normalized by the corresponding prescription dose. We evaluated OAR protection using lung D_{mean} , spinal cord D_{max} and $D_{0.03\text{cm}^3}$, heart D_{mean} , and liver D_{mean} . The volumetric-based constraints including lung $V_{5\text{Gy[RBE]}}$ and $V_{20\text{Gy[RBE]}}$, heart $V_{20\text{Gy[RBE]}}$, $V_{30\text{Gy[RBE]}}$ and $V_{40\text{Gy[RBE]}}$, kidney $V_{18\text{Gy[RBE]}}$, and liver $V_{30\text{Gy[RBE]}}$ were also calculated. Additionally, the absolute volume of stomach $V_{45\text{Gy[RBE]}}$ by cm³ was calculated.

2.F | Robustness quantification

Patient set up uncertainty is considered to be random and can be modelled as a Gaussian distribution. Range uncertainty is considered to be systematic, but range uncertainty of a large patient population can also be considered to be a Gaussian distribution.⁵⁰ The value of 3 mm setup uncertainty and 3% range uncertainty (equal to twice the standard deviation of the setup and range uncertainty distribution) is well-regarded in the proton therapy community. This is applicable for the treatment of distal esophageal carcinoma with plan robustness analysis and also in the use of image-guided radiation therapy.

For IMPT plans, 13 scenarios were taken into account, including 1 nominal and 12 perturbed scenarios representing uncertainty conditions. The range uncertainty due to the CT calibration error was assumed to be $\pm 3\%$ of the nominal beam range, and we rigidly shifted the isocenter of the patients in the anterior–posterior (A-P), superior–inferior (S-I), and right–left (R-L) directions by ± 3 mm respectively. The range and isocenter shift yielded 12 perturbed scenarios.

For VMAT plans, seven scenarios were considered, including one nominal scenario and six perturbed scenarios. The six perturbed scenarios were created by rigidly shifting the isocenter in the same canonical directions by a distance of 3 mm, but with no range uncertainty considerations. A DVH curve was generated for each uncertainty scenario and consequently a DVH band was formed corresponding to multiple uncertainty scenarios (Fig. 1). In order to evaluate the robustness of VMAT and IMPT treatment plans, the width of the DVH band was used as a surrogate for robustness indications. The width of the DVH band was the difference between the maximum and minimum of certain DVH indices (Fig. 1). The CTV-related DVH band width was normalized by the corresponding prescription dose. A smaller DVH band width meant better plan robustness. The DVH data for uncertainty analysis was exported from Eclipse, and the DVH band width was calculated using in-house developed software.

All treatment plans were designed to have CTV_{high} and CTV_{low} $D_{95\%}$ reaching at least 95% of the prescription doses in the worst-case scenarios for all CT phases.

2.G | Interplay effect evaluation

Dose evaluation software was developed to access the plans' ability to retain dose volume objectives under the influence of interplay effects. The dose evaluation software used the following machine parameters to model the time structure of spot delivery: spot delivery time per MU (i.e., dose rate), allowable extraction time (i.e., spill length), time for proton acceleration, deceleration and extraction setup, and time interval between consequent spots within the same energy layer (i.e., spot interval length). The relevant time sequence of spot delivery was shown in Fig. 2. Patient respiratory motion was modeled using 4D CTs and real-time position management (RPM) data using the Varian RPM system (Varian Medical Systems, Palo Alto, California). Every spot of each field per fraction was assigned to the corresponding respiratory phases according to their temporal relationship with the spot delivery sequence and patient-specific respiratory motion. The spot doses were summed to a reference phase using in-house developed deformable image registration

software.^{45,51} The starting phase of each field for each fraction was randomized to minimize the influence of starting phases. The interplay effects were evaluated using the fraction number in the prescription of IMPT treatment (Table S2).

The proton absolute dose calibration process follows IAEA TRS-398 protocol. The absolute dose of 1MU for a specific selected beam is 1 cGy. The beam used for calibration is: range 20 cm, spread out Bragg peak (SOBP) 10 cm, and field size 10 cm × 10 cm. It is composed of 27 energy layers (120–173.6 MeV), 289 spots in each layer with 6 mm spot spacing, a total of 200 MU. In IMPT treatment planning, we also used the iso-layer repainting to mitigate interplay effects.^{45,51,52} For respiratory motion less than or equal to 5 mm, the minimum and maximum MU limits in the proton machine were 0.003 and 0.04 MU, respectively. A smaller maximum MU limit of 0.01 MU was used for cases with respiratory motion greater than 5 mm. The purpose of a smaller maximum MU limit was to make the delivery system perform a larger number of iso-layer repainting, which in turn mitigated the impact of interplay effects.^{52,53} During the delivery process, if the intensity of one spot was larger than the maximum MU limit, the spot was split into multiple spots, and the split spots were appended in the spot list of the same energy layer and delivered individually. For the spots with MU less than the minimum MU limit: if the intensity of one spot was larger than half of the minimum MU limit (0.0015 MU), it was rounded up; and vice versa. For example, if the amplitude of breathing motion was smaller than 5 mm, a spot of 0.121 MU was split into three spots of 0.04 MU, and the residual 0.001 MU was neglected due to its value less than half of the minimum MU limit (0.0015 MU); a spot of 0.082 MU was be split into two spots of 0.04 MU and one spot of 0.003, since the remaining 0.002 MU was larger than 0.0015 MU. If the amplitude of breathing motion was larger than 5 mm, a spot of 0.121 MU was split into 12 spots of 0.01 MU, and the residual 0.001 MU was neglected due to its value less than half of the minimum MU limit (0.0015 MU); a spot of 0.082 MU was split into eight spots of 0.01 MU and one spot of 0.003, since the remaining 0.002 MU was larger than 0.0015 MU. No interplay effects were evaluated for VMAT planning, as it has been shown that the impact was small and not clinically significant.^{54,55}

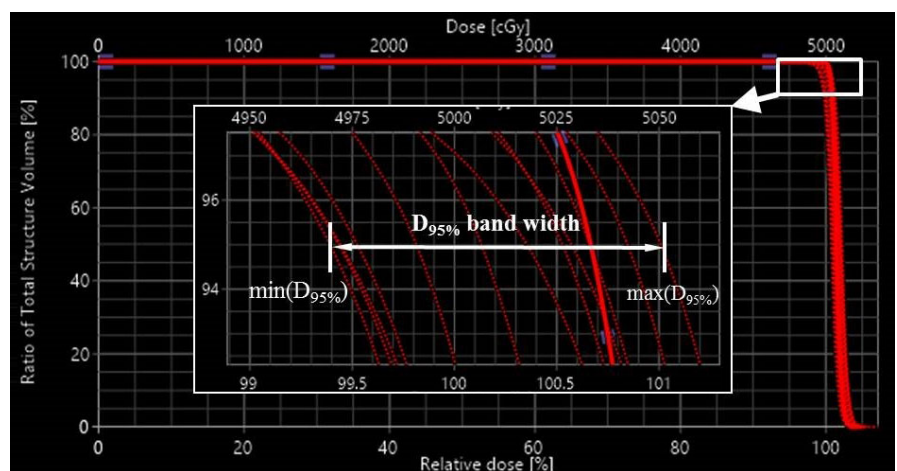


FIG. 1. A graphical representation of the dose-volume-histogram band width calculating method.

2.H | Statistical analysis

For fair comparison, all the treatment plans were normalized to have the same $CTV_{high} D_{95\%}$ in the nominal scenario relative to the prescription dose. Wilcoxon Rank-sum and Fisher's exact test were utilized to compare clinical characteristics between treatment populations for continuous and categorical data respectively. $P < 0.003$ ($0.05/16 = 0.003$) were considered statistically significant after adjusting for 16 different DVH indices. The results of the DVH indices from all patients were presented using Box-and-Whisker plotting which showed median values along with error bars. The maximum and minimum outlier points were located outside of the error bars which corresponded to 1.5 times the interquartile range above the upper and below the lower quartiles.

3 | RESULTS

3.A | Plan quality

Considering plan quality, we first compared the DVH indices of IMPT and VMAT in the nominal scenario (without any uncertainties considered). The IMPT plans performed significantly better in terms of liver $V_{30Gy[RBE]}$, lung D_{mean} and $V_{5Gy[RBE]}$, heart D_{mean} and $V_{20Gy[RBE]}$, and liver D_{mean} [Figs. 3(c)–3(f); Table 3]. Compared to the VMAT plans, IMPT plans did not significantly differ with respect to the following: $CTV_{high} D_{2cm^3}$ (normalized by the prescription doses), $CTV_{high} D_{5\%–D_{95\%}}$ (normalized by the prescription doses), CTV_{low} dose coverage, and protection of other OARs including kidney $V_{18Gy[RBE]}$, stomach $V_{45Gy[RBE]}$, cord D_{max} and $D_{0.03cm^3}$, lung $V_{20Gy[RBE]}$, and heart $V_{30Gy[RBE]}$ and $V_{40Gy[RBE]}$ [Figs. 3(a)–3(f); Table 3].

Some outliers were observed. For example, heart D_{mean} , $V_{20Gy[RBE]}$, and $V_{30Gy[RBE]}$ in IMPT [Figs. 3(e)–3(f)]. These outliers were found to come from the same patient. The heart was in close proximity to the CTV_{high} [Fig. S1(a)], which resulted in higher dose to the heart [Fig. S1(b)]. Similarly, close proximity of the heart to a tumor resulted in the outliers in heart $V_{30Gy[RBE]}$ and $V_{40Gy[RBE]}$ [Fig. 3(f)] in VMAT. To avoid vital organs including cord and lungs for this patient, we had to limit target dose coverage and homogeneity to some degree in VMAT, resulting in additional outliers in $D_{5\%–D_{95\%}}$ [Fig. 3(b)].

3.B | Plan robustness

Considering the plan robustness, the DVH index range of CTVs and OARs for both treatment groups were compared (Table 4). The robustness of IMPT plans was comparable to that of VMAT plans for kidney $V_{18Gy[RBE]}$, liver $V_{30Gy[RBE]}$, stomach $V_{45Gy[RBE]}$, lung D_{mean} , $V_{5Gy[RBE]}$ and $V_{20Gy[RBE]}$, cord D_{max} and $D_{0.03cm^3}$, liver D_{mean} , heart $V_{20Gy[RBE]}$ and $V_{30Gy[RBE]}$. The robustness of IMPT plans was statistically worse than that of VMAT plans for $CTV_{high} D_{95\%}$ (normalized by the prescription doses), $CTV_{high} D_{2cm^3}$ (normalized by the prescription doses), $CTV_{high} D_{5\%–D_{95\%}}$ (normalized by the prescription doses), $CTV_{low} D_{95\%}$ (normalized by the prescription doses), heart D_{mean} and $V_{40Gy[RBE]}$ (Fig. 4).

3.C | Interplay effect

Interplay effects were considered for all of the IMPT plans [see Fig. 5(a)–5(d)]. The median values of $CTV_{high} D_{95\%}$, D_{2cm^3} , and $D_{5\%–D_{95\%}}$ (normalized by the prescription doses) were 0.98, 1.06, and 0.06, respectively. The median value of $CTV_{low} D_{95\%}$ was 0.99.

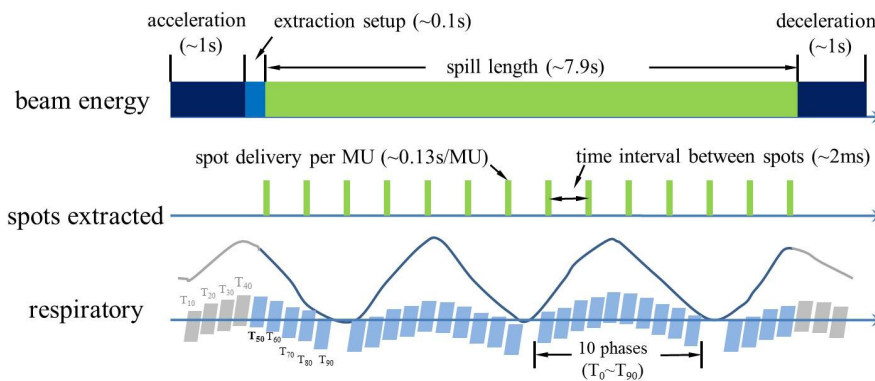
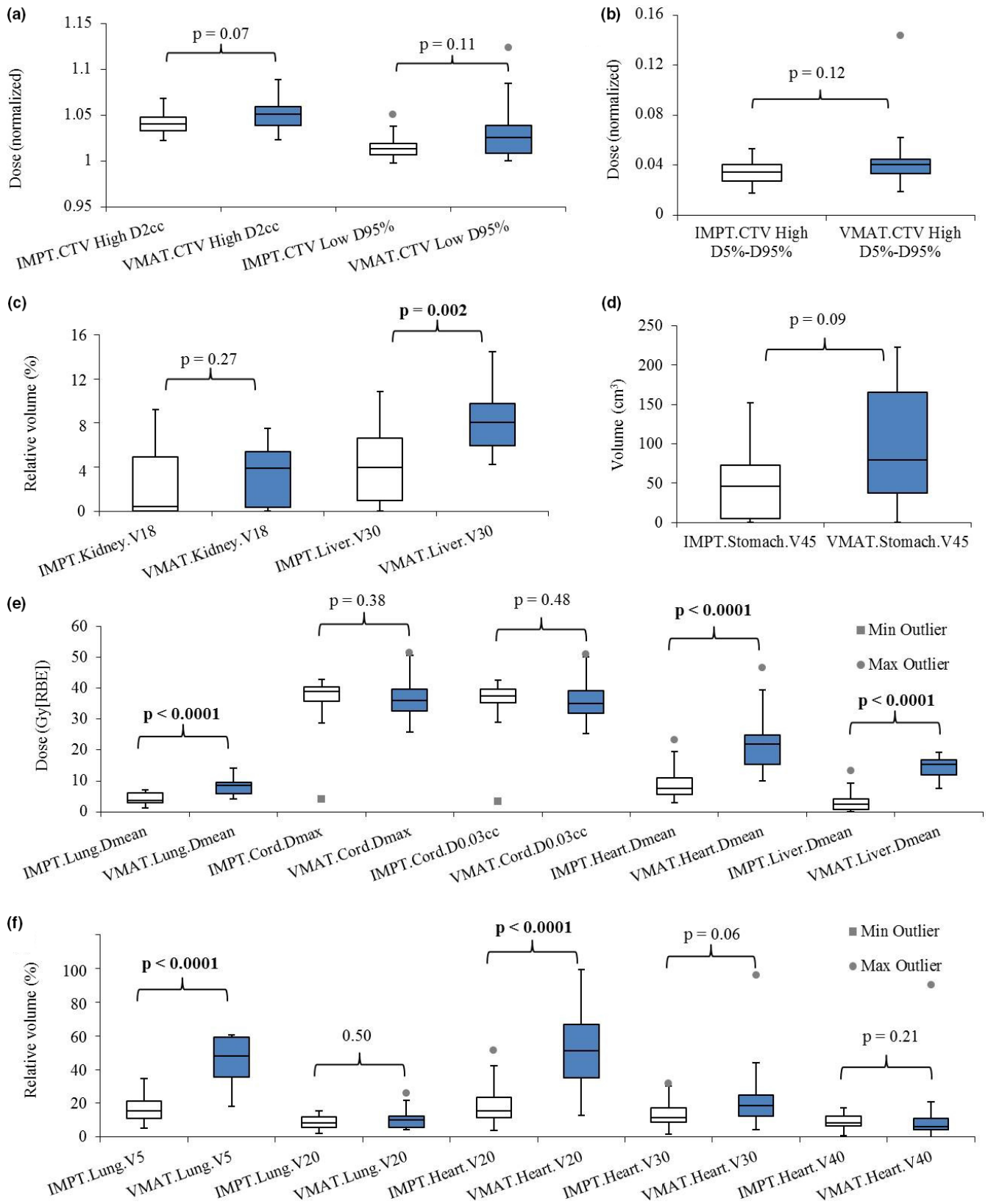


FIG. 2. Time dependent beam spot delivery sequence and modelling of interplay effects. A random starting phase (for example T_{50}) was used for each field per fraction to minimize the impact of the starting phase in the evaluation of interplay effects.

FIG. 3. Comparison of the dose-volume-histogram (DVH) indices between intensity-modulated proton therapy (IMPT) and volumetric-modulated arc therapy (VMAT) treatment plans in the nominal scenario. (a) clinical target volume ($CTV_{high} D_{2cm^3}$ and $CTV_{low} D_{95\%}$ normalized to the prescription dose. (b) $CTV_{high} D_{5\%–D_{95\%}}$ normalized to the prescription dose. (c) Kidney $V_{18Gy[RBE]}$ and $V_{30Gy[RBE]}$ in relative volume. (d) Absolute volume stomach $V_{45Gy[RBE]}$ in absolute volume. (e) Lung D_{mean} , cord D_{max} and $D_{0.03cm^3}$, heart D_{mean} , and liver D_{mean} in absolute dose. (f) Lung $V_{5Gy[RBE]}$ and $V_{20Gy[RBE]}$, heart $V_{20Gy[RBE]}$, $V_{30Gy[RBE]}$ and $V_{40Gy[RBE]}$ in relative volume. In each box plot, the five horizontal lines from top to bottom are the maximum, third quartile, median, first quartile, and minimum value of the corresponding DVH index for the whole group excluding the outliers respectively. The grey points are the outliers as defined in Statistical Analysis subsection. Numbers at the top of the columns are P -values from Wilcoxon rank sum testing. The blue boxes are the IMPT results and white boxes are the VMAT results. Abbreviations: RBE = relative biological effectiveness.



Median values of total lung $V_{5\text{Gy[RBE]}}$ and $V_{20\text{Gy[RBE]}}$, heart $V_{20\text{Gy[RBE]}}$, $V_{30\text{Gy[RBE]}}$ and $V_{40\text{Gy[RBE]}}$, liver $V_{30\text{Gy[RBE]}}$, left, right and total kidney $V_{18\text{Gy[RBE]}}$ were 17.4%, 8.7%, 16.1%, 11.6%, 7.7%, 3.8%, 0%, 0%,

and 0.09% respectively. Median values of total lung D_{mean} , spinal cord D_{max} and $D_{0.03\text{cm}^3}$, heart D_{mean} , liver D_{mean} were 4.0 Gy[RBE], 39.2 Gy[RBE], 37.8 Gy[RBE], 7.8 Gy[RBE], 2.7 Gy[RBE] respectively.

TABLE 3 The comparison of plan quality using dose-volume-histogram (DVH) indices.

DVH index	VMAT	IMPT	P-value
CTV _{high} D _{2cm³} (normalized)	105%	104%	0.071
CTV _{high} D _{5%–D_{95%}} (normalized)	4.0%	3.5%	0.12
CTV _{low} D _{95%} (normalized)	103%	101%	0.11
Liver V _{30Gy[RBE]} (%)	8.05%	3.97%	0.0023
Liver D _{mean} (Gy[RBE])	15.27	2.58	<0.0001
Lung D _{mean} (Gy[RBE])	8.55	3.73	<0.0001
Lung V _{5Gy[RBE]} (%)	48.11%	15.61%	<0.0001
Lung V _{20Gy[RBE]} (%)	10.32%	8.57%	0.50
Heart D _{mean} (Gy[RBE])	21.94	7.63	<0.0001
Heart V _{20Gy[RBE]} (%)	51.06%	15.61%	<0.0001
Heart V _{30Gy[RBE]} (%)	18.52%	11.48%	0.061
Heart V _{40Gy[RBE]} (%)	8.38%	6.31%	0.21
Cord D _{max} (Gy[RBE])	36.00	38.92	0.38
Cord D _{0.03cm³} (Gy[RBE])	35.05	37.49	0.48
Kidney V _{18Gy[RBE]} (%)	3.88%	0.41%	0.27
Stomach V _{45Gy[RBE]} (cm ³)	79.92	45.71	0.093

CTV, clinical target volume; IMPT, intensity-modulated proton therapy; VMAT, volumetric-modulated arc therapy.

Bold values represent significant difference between IMPT and VMAT DVH indices.

4 | DISCUSSION

The purpose of this comparative planning study was to evaluate the process and feasibility of small-spot machine IMPT in the treatment for distal esophagus carcinoma as well as provide a dosimetric comparison to VMAT in regard to plan quality and robustness.

Small spot-size IMPT is an attractive modality for the treatment of esophageal cancer. Our study compared IMPT with spot sizes of 2–6 mm (σ) with VMAT. We found that small-spot IMPT achieved similar plan quality as VMAT in terms of target dose coverage, homogeneity, and sparing of most OARs. More importantly, it significantly lowered heart, liver, and bilateral lung doses as compared to VMAT. As a result, IMPT will likely reduce the incidence and also severity of RT-induced cardiac and pulmonary toxicities in the long-term and perioperatively. However, the equipoise of such considerations and clinically meaningful significance over conventional treatments such as IMRT or VMAT remain undefined for distal esophageal cancer therapies.

TABLE 4 The comparison of plan robustness using the width of dose-volume-histogram (DVH) index bands.

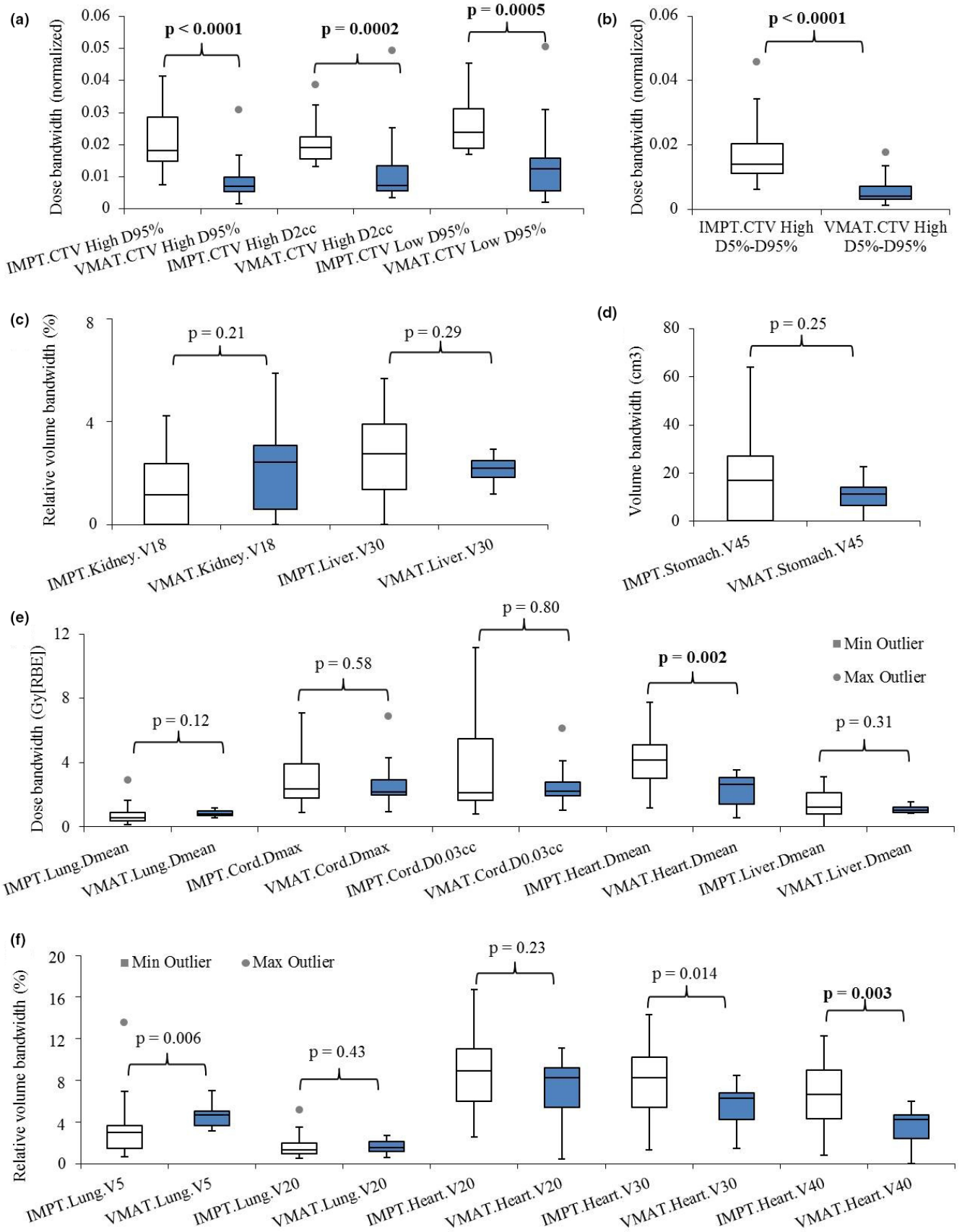
DVH index	VMAT	IMPT	P-value
CTV _{high} D _{95%} (normalized)	0.7%	1.8%	<0.0001
CTV _{high} D _{2cm³} (normalized)	0.73%	1.9%	0.0002
CTV _{high} D _{5%–D_{95%}} (normalized)	0.41%	1.4%	<0.0001
CTV _{low} D _{95%} (normalized)	1.3%	2.4%	0.0005
Liver V _{30Gy[RBE]} (%)	2.19%	2.76%	0.29
Liver D _{mean} (Gy[RBE])	1.03	1.22	0.31
Lung D _{mean} (Gy[RBE])	0.80	0.58	0.12
Lung V _{5Gy[RBE]} (%)	4.73%	2.99%	0.0057
Lung V _{20Gy[RBE]} (%)	1.58%	1.31%	0.43
Heart D _{mean} (Gy[RBE])	2.65	4.13	0.002
Heart V _{20Gy[RBE]} (%)	8.28%	8.91%	0.23
Heart V _{30Gy[RBE]} (%)	6.31%	8.24%	0.014
Heart V _{40Gy[RBE]} (%)	4.28%	6.69%	0.0027
Cord D _{max} (Gy[RBE])	2.17	2.35	0.58
Cord D _{0.03cm³} (Gy[RBE])	2.21	2.12	0.80
Kidney V _{18Gy[RBE]} (%)	2.42%	1.16%	0.21
Stomach V _{45Gy[RBE]} (cm ³)	11.34	16.99	0.25

CTV, clinical target volume; IMPT, intensity-modulated proton therapy; VMAT, volumetric-modulated arc therapy.

Bold values represent significant difference between IMPT and VMAT DVH indices.

Although small-spot IMPT is potentially capable of producing better plans due to sharper penumbra, the uncertainties due to proton range and patient setup can greatly compromise plan qualities.^{56,57} Therefore, it is important to take into account the uncertainties in treatment planning when using a small-spot IMPT machine to treat patients with esophageal carcinoma. Our study demonstrates that small-spot IMPT treatment plans achieved clinically defined planning requirements in terms of plan robustness and met the clinical standards for RT. Small-spot IMPT should be considered for the routine treatment in patients with distal esophageal carcinoma. However, robustness relative to internal organ motion remains a major challenge in small-spot IMPT treatments for esophageal cancer.^{30,33,45,53} Compared with IMPT, VMAT has been shown to be more robust with respect to organ motions or anatomic changes.⁵⁸ Therefore, it is vital to consider and optimally mitigate the impact of respiratory motions when IMPT is clinically implemented. At our institution, we carefully consider respiratory motion

Fig. 4. Comparison of plan robustness between intensity-modulated proton therapy (IMPT) and volumetric-modulated arc therapy (VMAT) plans using the widths of the dose volume histogram (DVH) band pairs in the presence of uncertainties. (a) clinical target volume (CTV_{high}) D_{95%} and D_{2cm³}, and CTV_{low} D_{95%} normalized to the prescription dose. (b) CTV_{high} D_{5%–D_{95%}} normalized to the prescription dose. (c) Kidney V_{18Gy[RBE]} and liver V_{30Gy[RBE]} in relative volume. (d) Stomach V_{45Gy[RBE]} in absolute volume. (e) Lung D_{mean}, cord D_{max} and D_{0.03cm³}, heart D_{mean}, and liver D_{mean} in absolute dose. (f) Lung V_{5Gy[RBE]} and V_{20Gy[RBE]}, heart V_{20Gy[RBE]}, V_{30Gy[RBE]} and V_{40Gy[RBE]} in relative volume. In each box plot, the five horizontal lines from top to bottom are the maximum, third quartile, median, first quartile, and minimum value of the DVH band widths in the presence of uncertainties of the corresponding DVH index for the whole group excluding the outliers respectively. The grey points are the outliers as defined in Statistical Analysis subsection. Numbers at the top of the columns are P-values from Wilcoxon rank sum testing. The blue boxes are the IMPT results and white boxes are the VMAT results. Abbreviations: RBE = relative biological effectiveness.



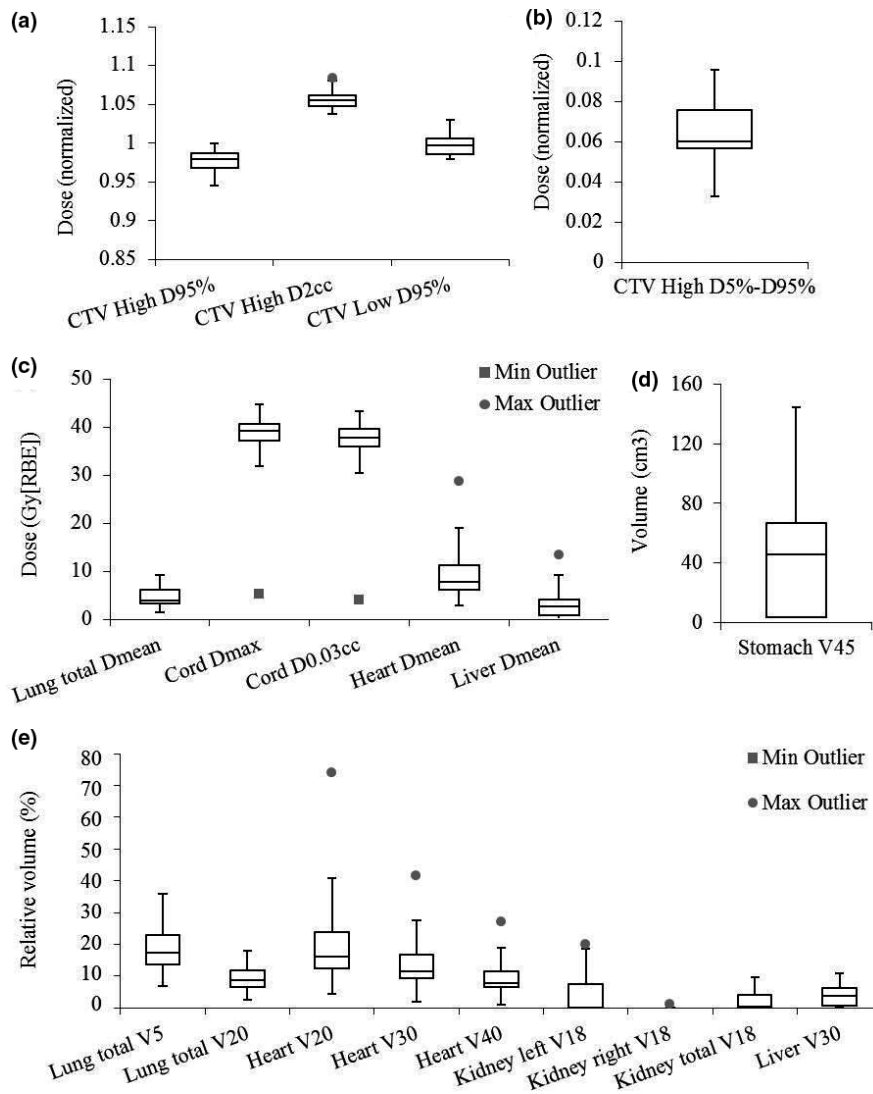


FIG. 5. Interplay effects evaluation in intensity-modulated proton therapy (IMPT) using the dose–volume histogram indices with interplay effects considered, including (a) clinical target volume (CTV_{high}) D_{95%} and D_{2cm³}, and CTV_{low} D_{95%} normalized to the prescription dose. (b) CTV_{high} D_{5%}–D_{95%} normalized to the prescription dose. (c) Lung D_{mean}, cord D_{max} and D_{0.03cm³}, heart D_{mean}, and liver D_{mean} in absolute dose. (d) Stomach V_{45Gy[RBE]} in absolute volume. (e) Lung V_{5Gy[RBE]} and V_{20Gy[RBE]}, heart V_{20Gy[RBE]}, V_{30Gy[RBE]} and V_{40Gy[RBE]}, left, right, and total kidneys V_{18Gy[RBE]}, and liver V_{30Gy[RBE]} in relative volume. In each box plot, the five horizontal lines from top to bottom are the maximum, third quartile, median, first quartile, and minimum value of the corresponding DVH index for the whole group excluding the outliers respectively. The grey points are the outliers as defined in Statistical Analysis subsection. The blue boxes are the IMPT results and white boxes are the VMAT results. At our clinics it is required to have CTV D_{95%} is at least 95% of the prescription dose for targets and the cord D_{max} < 45 Gy[RBE] under the influence of interplay effects. Abbreviations: RBE = relative biological effectiveness.

and its impact in our treatment planning guidelines. Compared with IMRT, large-spot IMPT using three beam angles (anteroposterior, right posterior oblique, and left posterior oblique) has been shown to deliver significantly lower lung D_{mean}, lung V_{5Gy[RBE]}, V_{10Gy[RBE]}, and V_{20Gy[RBE]}, and heart D_{mean} for distal esophageal cancer patients.³⁷ Compared with large-spot IMPT, small-spot IMPT is considered to have two benefits: (a) sharper penumbra, which results from smaller spot sizes, can improve the sparing of organs-at-risk and lead to lower radiation toxicities^{52,59}; (b) a larger number of spots, which are needed to cover the same tumors compared to large-spot IMPT, provides the TPS more freedom to compensate for the impact of uncertainties and interplay effects.⁵³

Most of the IMPT plans included in this study were generated using a SFO approach. The few remaining plans were generated using MFO with robust planning from our TPS.^{39–43} After the plan was optimized, we recalculated the dose distributions in the averaged 4D CT, maximum exhale phase CT, and maximum inhale phase CT without density override to evaluate the impact of respiratory motion. The original plan generated on the averaged CT was

adjusted until all dose distributions calculated on the averaged CT, maximum exhale phase CT, and maximum inhale phase CT without density override met the institutional dose constraints. All patients were observed to have both CTV_{high} and CTV_{low} D_{95%} at least 95% of the prescription dose with interplay effects taken into account. As a result, even though some of the patients included in this study have considerable amount of target motion as influenced by respiration (the median value of the amplitude of respiratory motions was 8 mm, with a range of 4 to 13 mm), all of our plans using small-spot IMPT still met the clinical requirement in terms of robustness and interplay effect considerations after these treatment planning measures were taken.

For distal esophageal carcinoma, we commonly used posterior oblique beams in our IMPT treatment: left–right oblique beams or superior–inferior oblique beams. Left–right oblique beams spare the spinal cord, but may be sensitive to respiratory motion (Fig. S2). Superior–inferior oblique beams spare the kidneys and have better plan robustness to respiratory motion since the beam does not travel through much of the diaphragm. The above beam angle

arrangements worked very well for most patients. However, in some scenarios, there were variants of beam arrangements for specific patients: if the patient was pretreated by radiation therapy and the dose in certain organs needed to be limited, three or four beams were used to spare certain organs (Table S2); if the patient had significant respiratory motion and two beams could not achieve the plan robustness of clinical requirements, then three or four beams were used to improve the plan robustness (Table S2).

This study has a number of limitations. First, the number of the patients included in this study was not sufficiently large nor were they matched. The results could be affected by interpatient variability and different planning skills. However, these represented consecutive samples of actual IMPT vs VMAT plans that were delivered in the clinic over a similar period of time. To address the aforementioned issues, a study with a larger patient population with both VMAT and IMPT plans generated for each patient is currently under way to further generalize our conclusions. Second, only a limited number of uncertainty scenarios were considered in this study, which might underestimate the impact of uncertainties in selected IMPT plans.

In the future, other tumor locations (cervical, proximal, and mid-esophageal) should be considered in further studies. Furthermore, more patient data with short/long-term clinical outcomes, perioperative complications, gradation-related toxicities, and patient-reported outcome should be reported to evaluate the potential clinical benefits of small-spot IMPT over VMAT plans.

5 | CONCLUSION

Compared to VMAT, small-spot IMPT significantly improves RT sparing of the heart, liver, and lungs, as well as achieves clinically acceptable plan robustness. The impact of interplay effects is small when proper treatment planning and respiratory motion measures are taken. Our results support the feasibility and acceptability for the routine clinical use of small-spot IMPT in patients with distal esophageal carcinoma.

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute (NCI) career developmental Award K25CA168984, Arizona Biomedical Research Commission Investigator Award, the Lawrence W. and Marilyn W. Matteson Fund for Cancer Research, and the Kemper Marley Foundation.

CONFLICTS OF INTEREST

None.

ETHICAL CONSIDERATIONS

This research was approved by the Mayo Clinic Arizona institutional review board (IRB, 13-005709). The informed consent was waived

by IRB protocol. Only image and dose-volume data were used in this study. All patient-related health information was removed for the study.

REFERENCES

- Herszenyi L, Tulassay Z. Epidemiology of gastrointestinal and liver tumors. *Eur Rev Med Pharmacol Sci*. 2010;14:249–258.
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA*. 2015;65:87–108.
- van Hagen P, Hulshof M, Van Lanschot J, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med*. 2012;366:2074–2084.
- Berger B, Belka C. Evidence-based radiation oncology: oesophagus. *Radiother Oncol*. 2009;92:276–290.
- Cooper JS, Guo MD, Hershkovic A, et al. Chemoradiotherapy of locally advanced esophageal cancer: long-term follow-up of a prospective randomized trial (RTOG 85–01). *JAMA*. 1999;281:1623–1627.
- Gharzai L, Verma V, Denniston KA, Bhurud AR, Bennion NR, Lin C. Radiation therapy and cardiac death in long-term survivors of esophageal cancer: an analysis of the Surveillance, Epidemiology, and End Result database. *PLoS ONE*. 2016;11:e0158916.
- Kole TP, Aghayere O, Kwah J, Yorke ED, Goodman KA. Comparison of heart and coronary artery doses associated with intensity-modulated radiotherapy versus three-dimensional conformal radiotherapy for distal esophageal cancer. *Int J Radiat Oncol Biol Phys*. 2012;83:1580–1586.
- Liu W, Patel SH, Shen JJ, et al. Robustness quantification methods comparison in volumetric modulated arc therapy to treat head and neck cancer. *Pract Radiat Oncol*. 2016;6:e269–e275.
- Van Benthuysen L, Hales L, Podgorsak MB. Volumetric modulated arc therapy vs. IMRT for the treatment of distal esophageal cancer. *Med Dosim*. 2012;36:404–409.
- Gong G, Wang R, Guo Y, et al. Reduced lung dose during radiotherapy for thoracic esophageal carcinoma: VMAT combined with active breathing control for moderate DIBH. *Radiat Oncol*. 2013;8:291.
- Deasy JO, Zuniga A, Apte A, Thorstad W, Bradley J. A new tumor control probability model fitted to head-and-neck and lung local control data and resulting implications for the impact of dose heterogeneity. *Int J Radiat Oncol Biol Phys*. 2013;87:S694.
- Register SP, Zhang X, Mohan R, Chang JY. Proton stereotactic body radiation therapy for clinically challenging cases of centrally and superiorly located stage I non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2011;80:1015–1022.
- Zhang X, Li Y, Pan X, et al. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage IIIB non-small-cell lung cancer: a virtual clinical study. *Int J Radiat Oncol Biol Phys*. 2010;77:357–366.
- Sejpal S, Komaki R, Tsao A, et al. Early findings on toxicity of proton beam therapy with concurrent chemotherapy for nonsmall cell lung cancer. *Cancer*. 2011;117:3004–3013.
- Chang JY, Li H, Zhu XR, et al. Clinical implementation of intensity modulated proton therapy for thoracic malignancies. *Int J Radiat Oncol Biol Phys*. 2014;90:809–818.
- Berman AT, Teo B-KK, Dolney D, et al. An in-silico comparison of proton beam and IMRT for postoperative radiotherapy in completely resected stage IIIA non-small cell lung cancer. *Radiat Oncol*. 2013;8:144.
- Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol*. 1996;41:111–124.

18. Kraus KM, Heath E, Oelfke U. Dosimetric consequences of tumor motion due to respiration for a scanned proton beam. *Phys Med Biol*. 2011;56:6563–6581.
19. Phillips MH, Pedroni E, Blattmann H, Boehringer T, Coray A, Scheib S. Effects of respiratory motion on dose uniformity with a charged-particle scanning method. *Phys Med Biol*. 1992;37:223–234.
20. Lambert J, Suchowerska N, McKenzie DR, Jackson M. Intrafractional motion during proton beam scanning. *Phys Med Biol*. 2005;50:4853–4862.
21. Grozinger SO, Bert C, Haberer T, Kraft G, Rietzel E. Motion compensation with a scanned ion beam: a technical feasibility study. *Radiat Oncol*. 2008;3:34.
22. Seco J, Robertson D, Trofimov A, Paganetti H. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. *Phys Med Biol*. 2009;54:N283–N294.
23. Grozinger SO, Rietzel E, Li Q, Bert C, Haberer T, Kraft G. Simulations to design an online motion compensation system for scanned particle beams. *Phys Med Biol*. 2006;51:3517–3531.
24. Dowdell S, Grassberger C, Sharp GC, Paganetti H. Interplay effects in proton scanning for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters. *Phys Med Biol*. 2013;58:4137–4156.
25. Grassberger C, Dowdell S, Lomax A, et al. Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. *Int J Radiat Oncol Biol Phys*. 2013;86:380–386.
26. Knopf A-C, Hong TS, Lomax A. Scanned proton radiotherapy for mobile targets—the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. *Phys Med Biol*. 2011;56:7257–7271.
27. Li Y, Kardar L, Li X, et al. On the interplay effects with proton scanning beams in stage III lung cancer. *Med Phys*. 2014;41:021721.
28. Kardar L, Li Y, Li X, et al. Evaluation and mitigation of the interplay effects of intensity modulated proton therapy for lung cancer in a clinical setting. *Pract Radiat Oncol*. 2014;4:e259–e268.
29. Bortfeld T, Jokivarsi K, Goitein M, Kung J, Jiang SB. Effects of intrafraction motion on IMRT dose delivery: statistical analysis and simulation. *Phys Med Biol*. 2002;47:2203–2220.
30. Liu W, Schild SE, Chang JY, et al. Exploratory study of 4D versus 3D robust optimization in intensity modulated proton therapy for lung cancer. *Int J Radiat Oncol Biol Phys*. 2016;95:523–533.
31. Moteabbed M, Yock TI, Depauw N, Madden TM, Kooy HM, Paganetti H. Impact of spot size and beam-shaping devices on the treatment plan quality for pencil beam scanning proton therapy. *Int J Radiat Oncol Biol Phys*. 2016;95:190–198.
32. Shirashi Y, Xu C, Yang J, Komaki R, Lin SH. Dosimetric comparison to the heart and cardiac substructure in a large cohort of esophageal cancer patients treated with proton beam therapy or Intensity-modulated radiation therapy. *Radiother Oncol*. 2017;125:48–54.
33. Chuong MD, Hallemeier CL, Jabbour SK, et al. Improving outcomes for esophageal cancer using proton beam therapy. *Int J Radiat Oncol Biol Phys*. 2016;95:488–497.
34. Lin SH, Komaki R, Liao Z, et al. Proton beam therapy and concurrent chemotherapy for esophageal cancer. *Int J Radiat Oncol Biol Phys*. 2012;83:e345–e351.
35. Zhang X, Guerrero TM, Mcguire SE, et al. Four-dimensional computed tomography-based treatment planning for intensity-modulated radiation therapy and proton therapy for distal esophageal cancer. *Int J Radiat Oncol Biol Phys*. 2008;72:278–287.
36. Lin C-Y, Huang W-Y, Jen Y-M, et al. Dosimetric and efficiency comparison of high-dose radiotherapy for esophageal cancer: volumetric modulated arc therapy versus fixed-field intensity-modulated radiotherapy. *Dis Esophagus*. 2014;27:585–590.
37. Welsh J, Gomez D, Palmer MB, et al. Intensity-modulated proton therapy further reduces normal tissue exposure during definitive therapy for locally advanced distal esophageal tumors: a dosimetric study. *Int J Radiat Oncol Biol Phys*. 2011;81:1336–1342.
38. Lin SH, Merrell KW, Shen J, et al. Multi-institutional analysis of radiation modality use and postoperative outcomes of neoadjuvant chemoradiation for esophageal cancer. *Radiother Oncol*. 2017;123:376–381.
39. Liu W, Frank SJ, Li X, et al. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. *Med Phys*. 2013;40:051711–051718.
40. Liu W, Frank SJ, Li X, Li Y, Zhu RX, Mohan R. PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization. *Med Phys*. 2013;40:021709–021708.
41. Liu W, Li Y, Li X, Cao W, Zhang X. Influence of robust optimization in intensity-modulated proton therapy with different dose delivery techniques. *Med Phys*. 2012;39:3089–3101.
42. Liu W, Mohan R, Park P, et al. Dosimetric benefits of robust treatment planning for intensity modulated proton therapy for base-of-skull cancers. *Pract Radiat Oncol*. 2014;4:384–391.
43. Liu W, Zhang X, Li Y, Mohan R. Robust optimization in intensity-modulated proton therapy. *Med Phys*. 2012;39:1079–1091.
44. Kang Y, Zhang X, Chang JY, et al. 4D proton treatment planning strategy for mobile lung tumors. *Int J Radiat Oncol Biol Phys*. 2007;67:906–914.
45. Liu W, Liao Z, Schild SE, et al. Impact of respiratory motion on worst-case scenario optimized intensity modulated proton therapy for lung cancers. *Pract Radiat Oncol*. 2015;5:e77–e86.
46. Wan Chan Tseung H, Ma J, Beltran C. A fast GPU-based Monte Carlo simulation of proton transport with detailed modeling of nonelastic interactions. *Med Phys*. 2015;42:2967–2978.
47. Shen J, Liu W, Stoker J, et al. An efficient method to determine double Gaussian fluence parameters in the eclipse™ proton pencil beam model. *Med Phys*. 2016;43:6544–6551.
48. Shen J, Lentz JM, Hu Y, et al. Using field size factors to characterize the in-air fluence of a proton machine with a range shifter. *Radiat Oncol*. 2017;12:52.
49. Shen J, Tryggestad E, Younkin JE, et al. Using experimentally determined proton spot scanning timing parameters to accurately model beam delivery time. *Med Phys*. 2017;44:5081–5088.
50. Pflugfelder D, Wilkens JJ, Oelfke U. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Phys Med Biol*. 2008;53:1689–1700.
51. Liu W, Schild S, Chang J, et al. A novel 4D robust optimization mitigates interplay effect in intensity-modulated proton therapy for lung cancer. *Med Phys*. 2015;42:3525.
52. Liu C, Sio TT, Deng W, et al. Small-spot intensity-modulated proton therapy and volumetric-modulated arc therapies for patients with locally advanced non-small-cell lung cancer: a dosimetric comparative study. *J Appl Clin Med Phys*. 2018;19:140–148.
53. Liu C, Schild SE, Chang JY, et al. Impact of spot size and spacing on the quality of robustly optimized intensity modulated proton therapy plans for lung cancer. *Int J Radiat Oncol Biol Phys*. 2018;101:479–489.
54. Hawkins M, Bedford J, Warrington A, Tait D. Volumetric modulated arc therapy planning for distal oesophageal malignancies. *Br J Radiol*. 2012;85:44–52.
55. Nicolini G, Ghosh-Laskar S, Shrivastava SK, et al. Volumetric modulation arc radiotherapy with flattening filter-free beams compared with static gantry IMRT and 3D conformal radiotherapy for advanced esophageal cancer: a feasibility study. *Int J Radiat Oncol Biol Phys*. 2012;84:553–560.
56. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Phys Med Biol*. 2008;53:1027–1042.

57. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Phys Med Biol*. 2008;53:1043–1056.
58. Chang JY, Jabbour SK, De Ruyscher D, et al. Consensus statement on proton therapy in early-stage and locally advanced non-small cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2016;95:505–516.
59. Yu NY, DeWees TA, Liu C, et al. Early outcomes of patients with locally advanced non-small cell lung cancer treated with intensity-modulated proton therapy (IMPT) vs. intensity-modulated radiation

therapy (IMRT): a single-institutional experience. *Adv Radiat Oncol*. 2019;104:234–235.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.