Advances in Radiation Oncology (2023) 8, 101292

advances in radiation oncology

www.advancesradonc.org

Scientific Article

Benchmarking Automated Machine Learning-Enhanced Planning With Ethos Against Manual and Knowledge-Based Planning for Locally Advanced Lung Cancer

Joel A. Pogue, PhD,[a,*] Carlos E. Cardenas, PhD,[a] Joseph Harms, PhD,[a] Michael H. Soike, MD,[a] Adam J. Kole, MD, PhD,[a] Craig S. Schneider, MD, PhD,[a] Christopher Veale, MD, MS,[a] Richard Popple, PhD,[a] Jean-Guy Belliveau, PhD,[a] Andrew M. McDonald, MD, MS,[a,b] and Dennis N. Stanley, PhD[a]

[a]*Department of Radiation Oncology, University of Alabama at Birmingham, Birmingham, Alabama; and* [b]*University of Alabama at Birmingham Institute for Cancer Outcomes and Survivorship, Birmingham, Alabama*

Received 10 April 2023; accepted 2 June 2023

Purpose: Currently, there is insufficient guidance for standard fractionation lung planning using the Varian Ethos adaptive treatment planning system and its unique intelligent optimization engine. Here, we address this gap in knowledge by developing a methodology to automatically generate high-quality Ethos treatment plans for locally advanced lung cancer.

Methods and Materials: Fifty patients previously treated with manually generated Eclipse plans for inoperable stage IIIA-IIIC non-small cell lung cancer were included in this institutional review board–approved retrospective study. Fifteen patient plans were used to iteratively optimize a planning template for the Daily Adaptive vs Non-Adaptive External Beam Radiation Therapy With Concurrent Chemotherapy for Locally Advanced Non-Small Cell Lung Cancer: A Prospective Randomized Trial of an Individualized Approach for Toxicity Reduction (ARTIA-Lung); the remaining 35 patients were automatically replanned without intervention. Ethos plan quality was benchmarked against clinical plans and reoptimized knowledge-based RapidPlan (RP) plans, then judged using standard dose-volume histogram metrics, adherence to clinical trial objectives, and qualitative review.

Results: Given equal prescription target coverage, Ethos-generated plans showed improved primary and nodal planning target volume V95% coverage ($P < .001$) and reduced lung gross tumor volume V5 Gy and esophagus D0.03 cc metrics ($P \leq .003$) but increased mean esophagus and brachial plexus D0.03 cc metrics ($P < .001$) compared with RP plans. Eighty percent, 49%, and 51% of Ethos, clinical, and RP plans, respectively, were "per protocol" or met "variation acceptable" ARTIA-Lung planning metrics. Three radiation oncologists qualitatively scored Ethos plans, and 78% of plans were clinically acceptable to all reviewing physicians, with no plans receiving scores requiring major changes.

Conclusions: A standard Ethos template produced lung radiation therapy plans with similar quality to RP plans, elucidating a viable approach for automated plan generation in the Ethos adaptive workspace.

Sources of support: This work had no specific funding.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Corresponding author: Joel A. Pogue, PhD; E-mail: japogue@uabmc.edu

https://doi.org/10.1016/j.adro.2023.101292

Introduction

Radiation therapy (RT) treatment planning remains a largely manual process in which dosimetrists, physicists, and physicians iteratively develop plans on a patient-by-patient basis. This planning approach uses significant departmental resources, such as time required to train personnel and develop quality treatment plans.[1] Additionally, manually generated plan quality is highly heterogeneous because of individual planner skill and experience.[2-4] Factors affecting plan quality include planner and radiation oncologist's choice of optimization dose and priority values, planning time limitations, and inconsistent use of optimization structures. Therefore, automated planning aims to decrease interplan variation and planning time while maintaining or improving plan quality.[5] This is achieved through standardization of structures of interest, optimization techniques, and priority values. Specific approaches include, but are not limited to, knowledge-based planning,[6-8] multicriteria optimization,[9-11] and template based planning.[12,13] To this end, the Ethos Adaptive RT platform (Varian Medical Systems, Palo Alto, CA) was designed with an automated treatment planning system (TPS) that generates plans from user-provided templates.[14]

Although the Ethos platform is relatively new, several aspects of plan generation have already been investigated. Mao et al[15] have shown that simulated adaptive lung therapy on the Ethos platform significantly improves target coverage and reduces normal tissue dose. Other studies have demonstrated that Ethos daily adaptive prostate and abdominal stereotactic body RT plans generally result in higher target coverage and a reduced risk of exceeding clinical organ-at-risk (OAR) thresholds compared with nonadapted plans.[16-18] The Ethos optimizer has been shown to generate similar or better quality pelvis and accelerated partial breast irradiation plans compared with manually generated Eclipse plans.[19,20] However, there has been no investigation of Ethos thoracic RT plan quality nor any analysis comparing automated planning in Ethos versus Eclipse. As the Ethos TPS introduces a unique intelligent optimization engine (IOE), and the optimizer drives RT plan generation, a quantitative and qualitative understanding of plan quality remains necessary. The primary endpoint of this retrospective treatment planning study was to develop an automated approach for creating high quality treatment plans in the Ethos adaptive workspace. As a secondary endpoint, Ethos-generated plans were quantitatively compared with the original clinical plans and RapidPlan (RP) plans from the same patients. Finally, Ethos-generated plans were qualitatively evaluated by experienced radiation oncologists.

Methods and Materials

Patient and volume description

Fifty patients previously treated at our institution for inoperable stage IIIA-IIIC non-small cell lung cancer between 2019 and 2022 were randomly selected for this prospective institutional review board (IRB-120703005) approved study. Patients were required to meet inclusion criteria for the prospective ARTIA-Lung clinical trial (Daily Adaptive vs Non-Adaptive External Beam Radiation Therapy With Concurrent Chemotherapy for Locally Advanced Non-Small Cell Lung Cancer: A Prospective Randomized Trial of an Individualized Approach for Toxicity Reduction (ARTIA-Lung); clinicaltrials.gov identifier NCT05488626). Accordingly, patients were excluded if they had contralateral hilar or supraclavicular lymph node involvement or distant metastases. Patients were immobilized according to institutional protocol and simulated head-first supine using a Phillips Brilliance big bore computed tomography (CT) scanner, if patient mobility allowed. Per institutional thoracic CT protocol, simulation CT scans extended from the inferior aspect of the cricoid through the entire liver using a slice thickness of 2 to 3 mm, and each patient received a respiratory-correlated 4-dimensional (4D) CT.

Targets and normal structures were generally delineated by the treating physician consistent with the Radiation Therapy Oncology Group 1106 study.[21] Positron emission tomography CT images were registered with the averaged 4D-CT to aid in delineation of primary and nodal gross tumor volumes (GTVp and GTVn, respectively), when available. Patients were treated free-breathing or with phase-based gating based on 4D motion assessment. Internal gross tumor volumes were defined as the union of GTVs drawn individually on each phase. Primary and nodal clinical target volumes (CTVp and CTVn, respectively) were typically generated by adding isotropic 5- or 7-mm margins to the internal gross tumor volumes, with manual cropping of the CTV at natural barriers to tumor invasion. Primary and nodal planning target volumes (PTVp and PTVn, respectively) were generated by adding margins between 5 and 10 mm to the CTVs, typically 5-mm isotropically (n = 38). Total PTV ranged from 95.1 to 1245.6 cm^3, with a mean value of 504.5 cm^3. All patients were prescribed 60 Gy in 30 fractions.

Treatment planning

Manually generated clinical Eclipse plans were initially calculated with the anisotropic analytical algorithm (AAA; version 13.6.23; Varian Medical Systems). Patients received intensity- modulated RT (IMRT) using 6 to 10

fields (n = 6) or volumetric-modulated arc therapy (VMAT) using 2 to 3 arcs (n = 44). To the contrary, the Ethos TPS calculates dose using AcurosXB (AXB) with heterogeneity correction on and dose to medium reporting mode selected (version 16.1.0; Varian Medical Systems). To account for systematic differences between AXB and AAA when calculating dose in the lungs,[22-24] all previously treated manual Eclipse plans were recalculated (preserving beam geometries and field weightings) using AXB with heterogeneity correction on and dose to medium reporting mode selected (version 15.5.11). Patients were treated using 6 MV flattened beams on a Varian Clinac 21IX and TrueBeam STX, each having a maximum dose rate of 600 MU/min and equipped with millennium 120 multileaf collimators (MLCs).[25] Additionally, each patient plan was reoptimized with AXB using our institution's primary thoracic Eclipse knowledge-based RP model, which uses the machine and MLCs originally used for patient treatment.

The Ethos TPS (version 1.1; Varian Medical Systems) is designed for cone beam CT-guided adaptive RT delivered on a Halcyon rotational linear accelerator, using a 6 MV flattening filter free beam with a maximum dose rate of 800 MU/min and employing double stacked MLC banks as its primary form of collimation.[26] For Ethos plan generation, all Eclipse clinical structure sets contoured before manual planning were anonymized and imported into the Ethos Treatment Management system (version 02.01.00; Varian Medical Systems). When replanning in Ethos, patients were randomly assigned to either the tuning (n = 15) or validation (n = 35) cohort until the tuning cohort had evenly distributed tumor laterality (left/right), ensuring the resulting template was robust to tumor location. Tuning cohort optimization priorities and dosimetric objectives were iteratively fine-tuned for maximum compliance with ARTIA-Lung clinical trial planning objectives, shown in Table 1. To investigate the quality of Ethos plans generated without planning bias (ie, trial-and-error planning), the remaining 35-patient validation cohort was automatically recalculated using the template resulting from the tuning cohort. The 35-patient validation cohort was deemed sufficiently large for TPS comparison, as paired, nonparametric analysis resulted in the ranking of 70 unique plans. Plans created using the template were evaluated as-is after the initial optimization (ie, no further planning).

Once each case was set up in the Ethos TPS and its planning template was approved, the IOE automatically optimized a plan for every user-selected, predefined beam geometry.[27] Ethos template planning objectives increase in priority with ascending order in the dose preview workspace, rather than by a cost function that varies with assigned priority number as in the Eclipse TPS. Six predefined geometries were selected for this work: equidistant 9- and 12-field IMRT plans, an ipsilateral 7-field IMRT plan, 2 and 3 full-arc VMAT plans, and a 2 half-arc VMAT plan. The reviewing physicist selected the optimal plan geometry for each patient according to the ARTIA-Lung metrics and hierarchy before comparison to clinical and RP plans. For final dose reporting, Ethos plans were exported and analyzed in the Eclipse TPS. All Ethos plans, clinical plans recalculated in AXB, and RP plans were normalized such that 95% of the total PTV received 60 Gy. Objective metrics and dose-volume histograms (DVH)

Table 1 ARTIA-Lung clinical trial planning objectives used to tune the Ethos template

Suggested priority	Structure	Dosimetric parameter	Per protocol	Variation acceptable
1	PTVp	V100% (%)	≥95%	≥90%
		V95% (%)	≥98%	≥95%
		D0.03 cc (%)	≤110%	≤115%
	PTVn	V100% (%)	≥95%	≥90%
		V95% (%)	≥98%	≥95%
		D0.03 cc (%)	≤110%	≤115%
	Spinal cord	D0.03 cc (Gy)	≤48 Gy	≤50 Gy
2	Lungs-GTV	V20 Gy (%)	≤33%	≤37%
		V5 Gy (%)	≤60%	≤65%
	Heart	V30 Gy (%)	≤50%	≤55%
		V45 Gy (%)	≤35%	≤40%
	Esophagus	D0.03 cc (Gy)	≤63 Gy	≤66 Gy
		Mean (Gy)	≤34 Gy	≤40 Gy
	Ipsilateral brachial plexus	D0.03 cc (Gy)	≤66 Gy	>66 Gy

Abbreviations: GTV = gross tumor volume; PTVn = nodal planning target volume; PTVp = primary planning target volume.

were extracted using the Eclipse Scripting Application Programming Interface for these plans.

The Wilcoxon paired, nonparametric test was used to test the difference between Ethos-generated plan metrics and clinical and RP plan metrics for all ARTIA-Lung planning objectives.[28] P values were calculated, without removal of outliers, using the SciPy library in Python (version 3). A Bonferroni correction was applied to adjust for multiple testing (n = 12); thus, $P \leq .004$ (.05/12) is considered significant when comparing dose metrics.

Qualitative evaluation

Three board certified radiation oncologists specializing in lung cancer treatment participated in Ethos plan reviews. All physicians initially reviewed 5 plans together to standardize the review process and normalize Likert scale plan scoring. Then, the physician reviewers were asked to objectively score each plan following our institution's clinical planning guidelines. Each physician then independently reviewed 30 of the remaining 45 plans, ensuring each of the 50 Ethos plans was reviewed by at least 2 physicians. The physicians were not provided additional clinical information regarding the cases and purely reviewed plan quality with anonymous patient identifiers. The physicians were not informed by physicists of the ARTIA-Lung clinical trial constraints, were not shown the template used for plan generation, and were not informed that plans had been calculated with AXB as opposed to AAA, which is the clinical standard. Instead, they judged plan quality according to their unique interpretation of the Likert scoring criteria outlined in Table 2.

Results

Generation of a planning template

A total of 129 Ethos initial intents and intent revisions were created throughout this analysis (94 for the tuning cohort, 35 for the validation cohort), resulting in the calculation and subsequent evaluation of 774 unique radiation plans. This volume of intents was necessary because many of the ARTIA-Lung planning objectives were not achieved for the tuning cohort when plans were directly optimized using the ARTIA-Lung objectives without using lower dose constraints and optimization structures. Thus, avoidance structures were created for any normal tissues potentially proximal to the target (lungs, heart, and esophagus). A 3-mm planning OAR spinal cord structure was also generated to aid in decreasing spinal cord dose and to account for patient setup uncertainty. The resulting template, found in Table 3, prioritized PTV coverage the most, followed by spinal cord avoidance then PTV hotspot reduction. Lung GTV (referred to as "lungs") V20 Gy was deemed the most important OAR metric aside from the spinal cord D0.03 cc; its avoidance structure was therefore placed at the bottom of priority 1. This template (right-sided) in XML format is included in the Supplementary Materials Table E1.

Dosimetric comparison

The Ethos validation cohort plans selected for comparison to clinical and RP plans included 9 equidistant 9-field plans (26%), 13 equidistant 12-field plans (37%), 11 ipsilateral 7-field plans (31%), 1 3-arc VMAT plan (3%), and 1 VMAT plan with 2 partial arcs (3%). The clinical and RP validation cohorts included 3 6-field IMRT plans, 2 7-field IMRT plans, 1 10-field IMRT plan, 14 VMAT plans with 2 partial arcs (ranging from 199°-224°), 3 VMAT plans with 3 partial arcs (ranging from 189°-199°), 6 2-arc VMAT plans, and 6 3-arc VMAT plans. Figure 1 shows boxplots comparing Ethos plan metrics to clinical and RP plan metrics for all ARTIA-Lung target and OAR objectives. Eighty percent (28/35), 49% (17/35), and 51% (18/35) of Ethos, clinical, and RP plans, respectively, were per protocol or met variation acceptable ARTIA-Lung metrics. For Ethos and RP plans, only lung V5 Gy metrics were variation unacceptable. The clinical cohort contained 2 PTVp D0.03 cc, 3 PTVn D0.03 cc, 15 lung V5

Table 2 Physician qualitative review grading scheme

Score	Description
5	Use as-is: Clinically acceptable plan that could be used for treatment without change.
4	Minor edits that are unnecessary: Reviewer prefers stylistic changes but considers current plan acceptable for treatment.
3	Minor edits that are necessary: Reviewer would require changes before treatment and the changes, in the judgment of the reviewer, can be implemented by minimal editing of the objectives.
2	Major edits: Reviewer would require changes before treatment, and the changes in the judgment of the reviewer would require significant modification of the objectives.
1	Unusable: The plan quality is so poor that it is deemed unsafe to deliver (ie, would likely result in harm to the patient).

Table 3 Ethos standard fractionation lung planning template

Priority	Structure	Structure derivation	Planning goal	Acceptable variation
1	PTVp		V100% ≥ 98.5%	V100% ≥ 90%
	PTVn		V100% ≥ 98.5%	V100% ≥ 90%
	PTVp		V95% ≥ 100%	V95% ≥ 95%
	PTVn		V95% ≥ 100%	V95% ≥ 95%
	Spinal cord		D0.03 cc ≤ 30 Gy	D0.03 cc ≤ 50 Gy
	PTVp		D0.03 cc ≤ 110%	D0.03 cc ≤ 115%
	PTVn		D0.03 cc ≤ 110%	D0.03 cc ≤ 115%
	_AvdLungs	Lungs (PTV + 0.3 cm)	V20 Gy ≤ 25%	V20 Gy ≤ 35%
2	SpinalCord_PRV03	Spinal cord + 0.3 cm	D0.03 cc ≤ 33 Gy	D0.03 cc ≤ 50 Gy
	_AvdLungs	Lungs (PTV + 0.3 cm)	D_{mean} ≤ 15 Gy	D_{mean} ≤ 20 Gy
	_AvdHeart	Heart (PTV + 0.3 cm)	V30 Gy ≤ 11%	V30 Gy ≤ 50%
	_AvdHeart	Heart (PTV + 0.3 cm)	V45 Gy ≤ 5%	V45 Gy ≤ 40%
	_AvdEsophagus	Esophagus (PTV + 0.3 cm)	D0.03 cc ≤ 57 Gy	D0.03 cc ≤ 63 Gy
	_AvdEsophagus	Esophagus (PTV + 0.3 cm)	D_{mean} ≤ 28 Gy	D_{mean} ≤ 34 Gy
	Lungs − GTV		V20 Gy ≤ 28%	V20 Gy ≤ 37%
	Lungs − GTV		D_{mean} ≤ 20 Gy	D_{mean} ≤ 22 Gy
	Heart		V30 Gy ≤ 12%	V30 Gy ≤ 55%
	Heart		V45 Gy ≤ 6%	V45 Gy ≤ 40%
	Esophagus		D0.03 cc ≤ 63 Gy	D0.03 cc ≤ 66 Gy
	Esophagus		D_{mean} ≤ 34 Gy	D_{mean} ≤ 40 Gy
	_AvdLungs	Lungs (PTV + 0.3 cm)	V5 Gy ≤ 55%	V5 Gy ≤ 60%
	Lungs − GTV		V5 Gy ≤ 60%	V5 Gy ≤ 65%
	Ipsilateral brachial plexus		D0.03 cc ≤ 66 Gy	D0.03 cc ≤ 66 Gy

Abbreviations: GTV = gross tumor volume; PTV = planning target volume; PTVn = nodal planning target volume; PTVp = primary planning target volume.
All optimization and planning organ-at-risk structures were generated in the Ethos workspace using the summarized derivations.

Gy, 1 esophagus D0.03 cc, and 1 brachial plexus D0.03 cc variation unacceptable metric. Ethos plans had higher mean esophagus dose than clinical plans (P = .002) and similar lung V20 Gy, esophagus D0.03 cc, and brachial plexus D0.03 cc metrics. Ethos showed significant improvement over clinical plans for all other metrics ($P \leq$ .001). Ethos-generated plans showed improved PTVp and PTVn V95% coverage ($P <$ .001) and reduced lung GTV V5 Gy and esophagus D0.03 cc metrics ($P \leq$ .003) but increased mean esophagus and brachial plexus D0.03 cc metrics ($P <$ .001) compared with RP plans.

Figure 2 shows validation cohort-averaged DVHs with standard deviation bounds for the PTVp, PTVn, lungs, heart, spinal cord, and esophagus. Ethos PTVp and PTVn have noticeably lower V105% values, and the mean Ethos DVH has sharper drop-off above prescription dose than clinical plans, indicating greater target homogeneity. Ethos spared the lungs below 10 Gy relative to clinical and RP plans. On average, RP resulted in increased heart sparing when considering heart doses below 30 Gy; in contrast, Ethos plans resulted in superior heart sparing for heart doses above 30 Gy. On average, RP plans have a lower spinal cord volume receiving below 28 Gy, but Ethos plans have a lower volume receiving above 28 Gy. Ethos plans generally produce less than 60 Gy outside of the target compared with clinical and RP plans.

Plan quality

As shown in Fig. 3, the mode Likert qualitative score for physicians A and C was 5, whereas physician B had a mode score of 4. Seventy-eight percent of plans were considered of clinically acceptable quality (score of 4 or 5) to each reviewing physician, whereas 22% (11/50) of plans had at least 1 physician score of 3 because of hotspots in OARs or balancing of PTV coverage to OAR sparing. No plans received a score of 1 or 2 from any physician. The average physician scores for the tuning and validation cohorts are 4.36 and 4.39, respectively, suggesting no

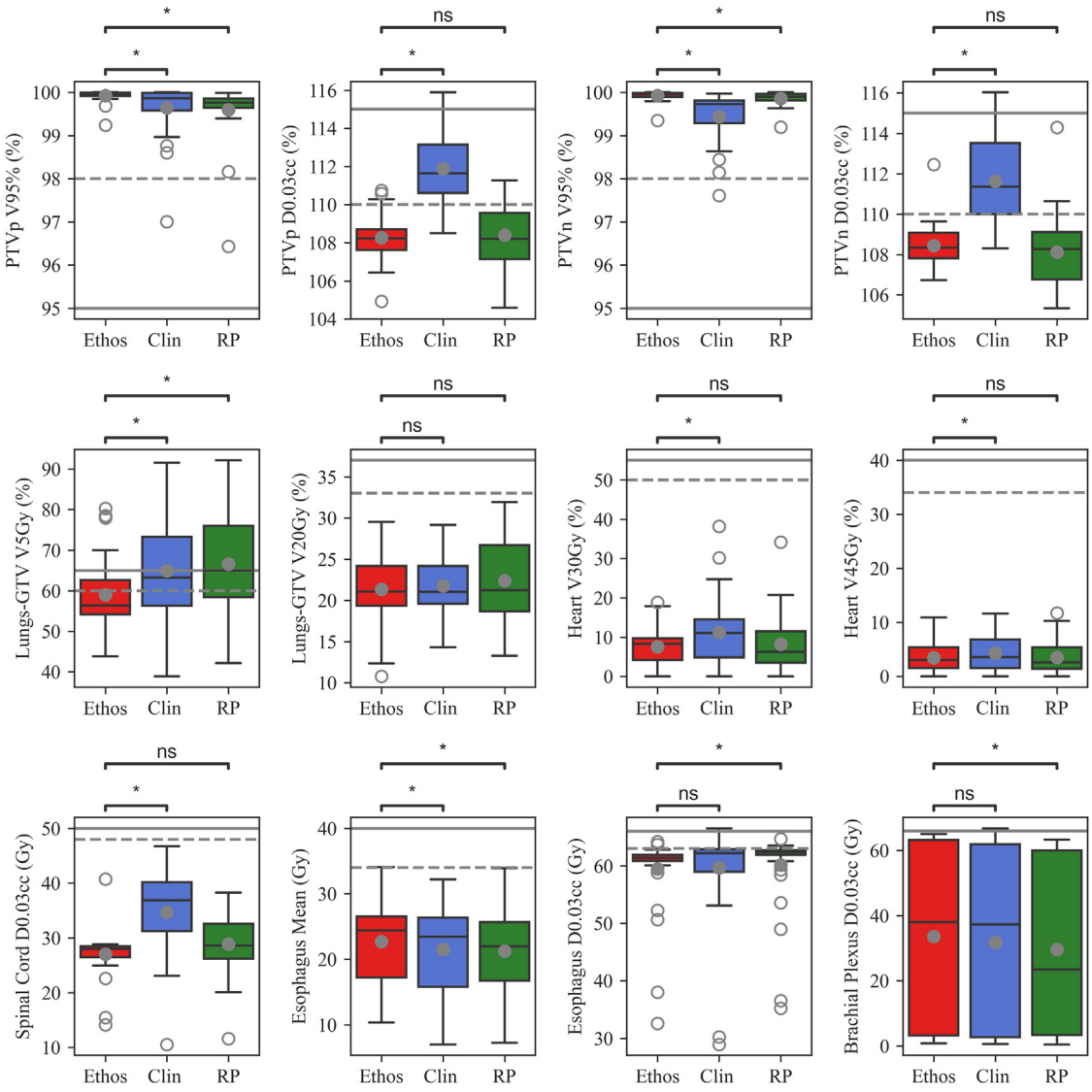

Figure 1 Boxplots summarizing Ethos, clinical plans recalculated in AcurosXB, and RapidPlan cohort ARTIA-Lung planning metrics. Open and closed circles indicate outlier and mean values, respectively. Dashed and solid lines indicate per protocol and variation acceptable constraints. Significance values obtained via the Wilcoxon signed rank test are annotated as follows: ns: $P > 0.004$; *: $P \leq 0.004$. *Abbreviations:* Clin = clinical; GTV = gross tumor volume; ns = not significant; PTV = planning target volume; RP = RapidPlan.

decrease in quality when plans were generated without bias.

Discussion

In this study we present, to our knowledge, the largest Ethos treatment planning study of a single treatment site and the only thoracic Ethos planning study, where automatically generated RT treatment plans for locally advanced lung cancer are quantitatively and qualitatively evaluated. Evaluation of the final planning template resulted in 80% of Ethos-generated plans meeting planning objectives without user input. These findings are supported by independent qualitative evaluation of individual plans, which showed that a large majority of plans (78%) for lung would be approved by multiple physicians as-is or with minor edits that are not necessary (ie, stylistic in nature and likely not clinically effective).

The high ratio of IMRT to VMAT plans in the validation cohort, along with the increase in time for VMAT dose calculation relative to IMRT, suggest that it may be clinically optimal to only perform IMRT calculations using this template, as adaptive treatment plans are calculated with patients on the treatment table. These results are in good agreement with Calmels et al,[19] who observed that, for most planning objectives, the Ethos TPS generates plans with similar or better target coverage and OAR

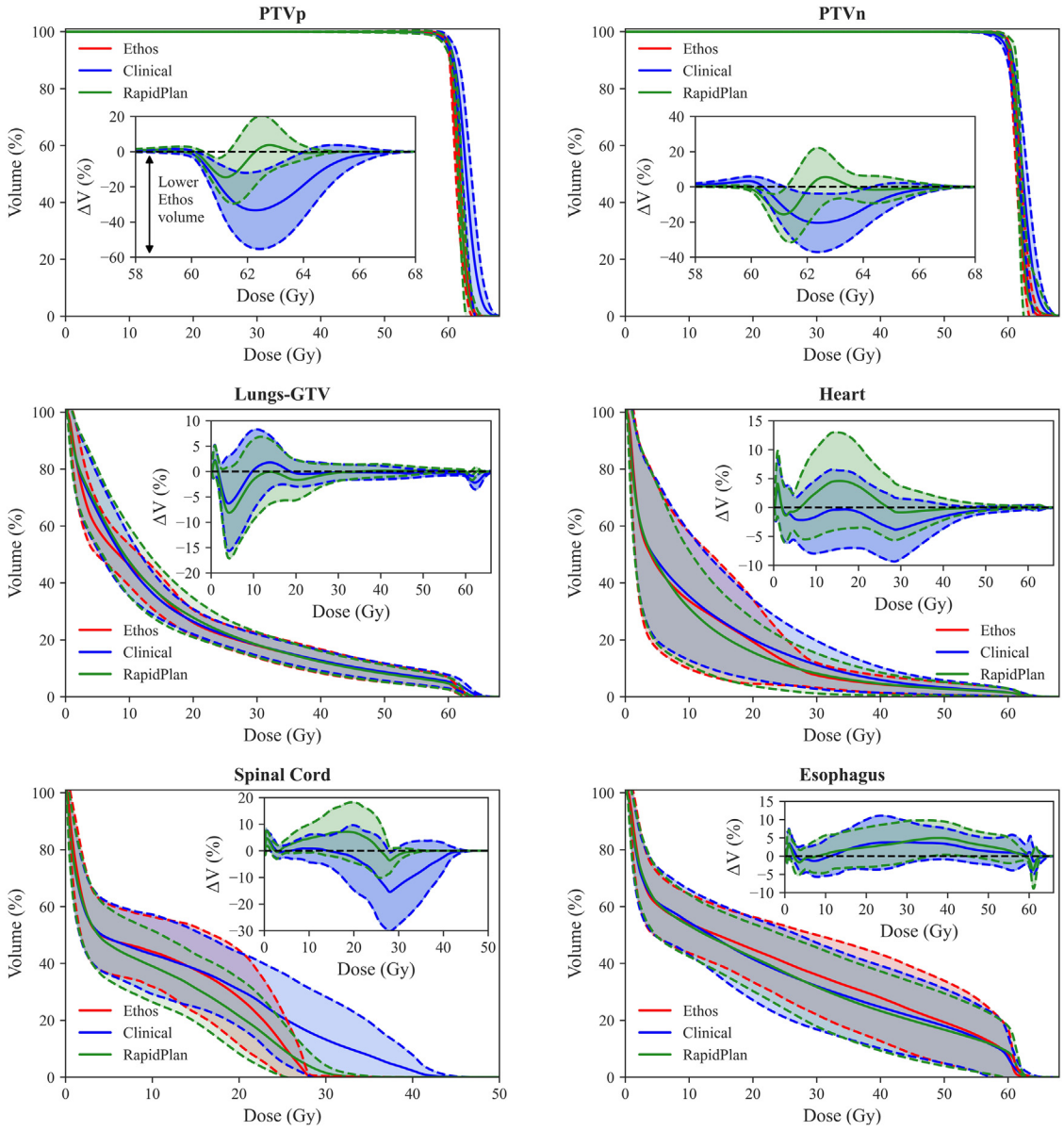

Figure 2 Validation cohort dose volume histogram comparison of Ethos, clinical plans recalculated in AcurosXB, and RapidPlan plans. Shaded areas illustrate the mean ± standard deviation and insets elucidate the difference between mean population dose volume histograms (ie, Ethos plan volume minus alternative plan volume). *Abbreviations:* GTV = gross tumor volume; PTV = planning target volume; PTVn = nodal planning target volume PTVp = primary planning target volume.

sparing compared with manually generated Eclipse VMAT plans for anal, rectal, and prostate cancers. It should be noted that further studies are needed to deconflate the effects of the Ethos treatment machine and optimizer on plan quality relative to the Eclipse TPS. More explicitly, a controlled experiment is required to determine what portion of Ethos plan difference is attributable to the double banked MLC and what portion is attributable to its IOE.

Although nonparametric testing is appropriate for analyzing differences in dose metric distributions that are not normal, nonparametric tests often fail to capture the magnitude of difference between 2 samples. As an example, the Wilcoxon signed rank test yields a significant improvement ($P < .001$) of Ethos PTVp V95% metrics over RP plans because there is a high probability that a randomly sampled patient will have a V95% value that is higher in Ethos versus RP. However, the means of the 2 V95% distributions are only 0.2% different. Thus, significant P values may not all translate to clinically meaningful improvements. Because of this, DVH curve comparison (Fig. 2) is used to elucidate TPS dose difference continuously as a function of target/OAR volume when necessary.

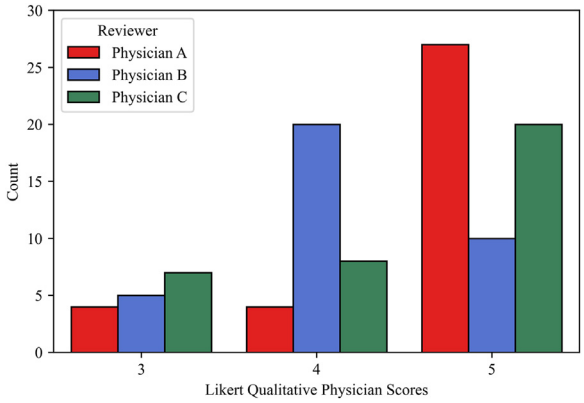

Figure 3 Likert qualitative plan review results by physician. The grading scale ranged from 1 to 5, but no scores below 3 were given.

Higher Ethos esophagus and brachial plexus metrics are due to the architecture of the optimizer in the dose preview workspace. Although all Eclipse TPS structures can be given an equal priority and therefore have a similar effect on the cost function, Ethos optimizer priorities are necessarily different from each other. Therefore, esophagus objectives would need to be moved higher than spinal cord, lung, or heart objectives in the dose preview workspace to further improve Ethos esophagus metrics relative to clinical and RP plans. However, the spinal cord, lungs, and heart were deemed more critical structures, and were therefore assigned a higher priority.

In this study, we recalculated clinical AAA plans in AXB and normalized to 95% prescription coverage. In AAA, none of the validation cohort PTVp and PTVn D0.03 cc metrics exceeded 115%, and none of the esophagus or brachial plexus D0.03 cc metrics exceeded 66 Gy. After recalculation in AXB, 5 validation cohort plans failed to achieve these constraints. This is consistent with literature that reports increased heterogeneity when transitioning from AAA to AXB.[22,23]

Although the original clinical intent cannot be known retrospectively, nor were the clinical or RP plans generated according to the ARTIA-Lung clinical trial, it is assumed that lung dose was limited to the extent possible during the initial planning process. The fact that significantly more Ethos plans met the lung V5 Gy objective suggests the Ethos template proposed in this work is well-tuned for limiting low levels of lung dose, thus complying with the ARTIA-Lung lung V5 Gy planning objective.

During physician plan review, disease extent and locale were considered to judge plan difficulty, and dose distributions were evaluated on a slice-by-slice basis, ensuring that the location of the dose inside the target or OAR was appropriate. The clinical effects of high dose "streaking" and hotspot location were also judged. Because reviews were unbiased, plan quality was based on reviewer-specific target and OAR objectives that were subjective and often more stringent than the trial objectives. Physicians were not satisfied with plans simply because all dose metrics were achieved but were looking instead for techniques that would allow plans to further improve. Generally speaking, physician A preferred target coverage over OAR sparing, physician B preferred minimizing maximum doses within OARs at the expense of target coverage, and physician C preferred that the target and OARs be somewhat equally prioritized. To that end, the goal of this work was to elucidate planning techniques for generating high quality, automated lung plans in the Ethos workspace, not to propose that certain target and OAR constraints be implemented in a particular clinic. Additionally, the goal of the ARTIA-Lung trial is to reduce target margins for adaptive treatment. With smaller target margins there will be less overlap between OARs and targets, and thus a lower dose to OARs.[29] This template could therefore potentially generate superior plans for patients treated adaptively.

Calmels et al[19] observed that median active plan preparation time ranged from 7.0 to 15.0 minutes for prostate, rectum, and anal cases. The median amount of passive times required for the IOE to generate plans ranged from 2.2 to 4.1 minutes for IMRT plans and 10.1 to 18.1 minutes for VMAT plans.[29] They then estimated that the time required for a planner to manually generate a plan is between 60 and 120 minutes. Although we did not document automated or manual planning times, our observations are consistent with the findings of Calmels et al.[19] We estimate that each lung IMRT and VMAT plan required 3 to 4 minutes of passive optimization time and approximately 15 minutes of passive optimization time, respectively; thus the 6 Ethos plans optimized for each patient in this work required approximately 1 hour to generate. However, unlike Calmels et al,[19] who altered failing planning objectives in the dose preview workspace before optimization, our work is entirely automated (ie, we used the same template for every patient in this study). Therefore, we estimate that only approximately 5 minutes of active planner time is required to choose the template and planning image, then associate the structure set. We further speculate that 1 to 2 hours is a good estimate for the active time required for a skilled planner to generate the clinical plans in this study.

Similar to the Ethos plans, RP plans require around 5 minutes to associate the structure set, set the beam geometry, and pair structures from the patient to structures in the model. After this initial setup, the RP plans used in this work were calculated automatically with no intervention as described by Harms et al,[30] with optimization and dose calculation together taking on average 15.0 $\pm$ 4.9 minutes.

To that end, automated planning in Ethos does not require additional interaction from physics once the planning templates are obtained, assuming the optimization objectives do not require altering. In the same way, dosimetrists can easily use RP without further physics intervention once the model has been built and validated.

Future work includes the clinical implementation and prospective clinical use of the resulting planning template. Furthermore, the methodology presented in this study will be translated to subsequent analyses focusing on extending our library of planning templates for other cancer sites. Another potential avenue for future study entails directly comparing Ethos halcyon plans generated using the Eclipse optimizer to those generated by the Ethos IOE.

The content and framework of this manuscript were constructed for consistency with the recently published RT treatment planning guidelines for generating high-quality planning studies.[31] The agreed-upon self-assessment score of 2 authors was 94% (197/210). The resulting spreadsheet is shared as Supplementary Materials Table E2.

Conclusion

The Ethos planning template developed in this study was applied to 50 patients with stage IIIA-IIIC non-small cell lung cancer. Automatically generated Ethos plans were overall similar in quality to RP plans, with per protocol or variation acceptable dosimetry in 80% of cases, without intervention or replanning. Seventy-eight percent of automatically generated plans were deemed clinically acceptable by multiple physicians. This template, or a variation thereof, enables an automated, high-quality planning approach for patients requiring adaptive standard fractionation treatment for locally advanced lung cancer.

Disclosures

Richard Popple reports that his institution, University of Alabama at Birmingham, has product evaluation agreements and research grants with Varian Medical Systems. He also reports that he has a patent licensed by UAB Research Foundation to Varian Medical Systems, has received honoraria for presentations on behalf of Varian Medical Systems, has received a stipend to speak at Sun Nuclear meetings, and that Varian Medical Systems provides equipment to UAB as a part of a product evaluation agreement. Andrew McDonald and his institution, University of Alabama at Birmingham, have received payment from Varian Medical Systems. Dennis N. Stanley and his institution, University of Alabama at Birmingham, have received payment from Varian Medical Systems.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.adro.2023.101292.

References

1. Winkel D, Bol GH, Van Asselen B, et al. Development and clinical introduction of automated radiotherapy treatment planning for prostate cancer. *Phys Med Biol*. 2016;61:8587-8595.
2. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract Radiat Oncol*. 2012;2:296-305.
3. Moore KL, Brame RS, Low DA, Mutic S. Quantitative metrics for assessing plan quality. *Semin Radiat Oncol*. 2012;22:62-69.
4. Batumalai V, Jameson MG, Forstner DF, Vial P, Holloway LC. How important is dosimetrist experience for intensity modulated radiation therapy? A comparative analysis of a head and neck case. *Pract Radiat Oncol*. 2013;3:e99-e106.
5. Moore KL, Appenzoller LM, Tan J, Michalski JM, Thorstad W, Mutic S. Clinical implementation of dose-volume histogram predictions for organs-at-risk in IMRT planning. *J Phys Conf Ser*. 2014;489: 012055.
6. Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. *Med Phys*. 2019;46:2760-2775.
7. Tambe NS, Pires IM, Moore C, Cawthorne C, Beavis AW. Validation of in-house knowledge-based planning model for advance-stage lung cancer patients treated using VMAT radiotherapy. *Br J Radiol*. 2020;93: 20190535.
8. Li N, Carmona R, Sirak I, et al. Highly efficient training, refinement, and validation of a knowledge-based planning quality-control system for radiation therapy clinical trials. *Int J Radiat Oncol*. 2017;97:164-172.
9. Naccarato S, Rigo M, Pellegrini R, et al. Automated planning for prostate stereotactic body radiation therapy on the 1.5 T MR-Linac. *Adv Radiat Oncol*. 2022;7: 100865.
10. Craft DL, Hong TS, Shih HA, Bortfeld TR. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. *Int J Radiat Oncol*. 2012;82:e83-e90.
11. Kierkels RG, Visser R, Bijl HP, et al. Multicriteria optimization enables less experienced planners to efficiently produce high quality treatment plans in head and neck cancer radiotherapy. *Radiat Oncol*. 2015;10:1-9.
12. Cilla S, Ianiro A, Romano C, et al. Template-based automation of treatment planning in advanced radiotherapy: A comprehensive dosimetric and clinical evaluation. *Sci Rep-UK*. 2020;10:1-13.
13. Vanderstraeten B, Goddeeris B, Vandecasteele K, Van Eijkeren M, De Wagter C, Lievens Y. Automated instead of manual treatment planning? A plan comparison based on dose-volume statistics and clinical preference. *Int J Radiat Oncol*. 2018;102:443-450.
14. Sibolt P, Andersson LM, Calmels L, et al. Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region. *Phys Imaging Radiat Oncol*. 2021;17:1-7.
15. Mao W, Riess J, Kim J, et al. Evaluation of auto-contouring and dose distributions for online adaptive radiation therapy of patients with locally advanced lung cancers. *Pract Radiat Oncol*. 2022;12:e329-e338.
16. Moazzezi M, Rose B, Kisling K, Moore KL, Ray X. Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited CBCT auto-segmentation. *J Appl Clin Med Phys*. 2021;22:82-93.
17. Byrne M, Archibald-Heeren B, Hu Y, et al. Varian ethos online adaptive radiotherapy for prostate cancer: Early results of contouring accuracy, treatment plan quality, and treatment time. *J Appl Clin Med Phys*. 2022;23:e13479.
18. Schiff JP, Stowe HB, Price A, et al. In silico trial of Computed Tomography-Guided Stereotactic Adaptive Radiotherapy (CT-STAR) for the treatment of abdominal oligometastases. *Int J Radiat Oncol*. 2022;114:1022-1031.

19. Calmels L, Sibolt P, Åström LM, et al. Evaluation of an automated template-based treatment planning system for radiotherapy of anal, rectal and prostate cancer. *Tech Innov Patient Support Radiat Oncol*. 2022;22:30-36.
20. Pogue JA, Cardenas CE, Cao Y, et al. Leveraging intelligent optimization for automated, cardiac-sparing accelerated partial breast treatment planning. *Front Oncol*. 2023;13: 1130119.
21. Ritter T, Quint DJ, Senan S, et al. Consideration of dose limits for organs at risk of thoracic radiotherapy: Atlas for lung, proximal bronchial tree, esophagus, spinal cord, ribs, and brachial plexus. *Int J Radiat Oncol*. 2011;81:1442-1457.
22. Fleming C, O'Keeffe S, McDermott R, Dunne M, McClean B, Vintró LL. The influence of Acuros XB on dose volume histogram metrics and tumour control probability modelling in locally advanced non-small cell lung cancer. *Phys Medica*. 2021;81:295-301.
23. Fogliata A, Nicolini G, Clivio A, Vanetti E, Cozzi L. Critical appraisal of Acuros XB and anisotropic analytic algorithm dose calculation in advanced non-small-cell lung cancer treatments. *Int J Radiat Oncol*. 2012;83:1587-1595.
24. Zifodya JM, Challens CH, Hsieh WL. From AAA to Acuros XB-clinical implications of selecting either Acuros XB dose-to-water or dose-to-medium. *Australas Phys Eng S*. 2016;39:431-439.
25. Chang Z, Wu Q, Adamson J, et al. Commissioning and dosimetric characteristics of TrueBeam system: Composite data of three TrueBeam machines. *Med Phys*. 2012;39:6981-7018.
26. Lim TY, Dragojević I, Hoffman D, Flores-Martinez E, Kim GY. Characterization of the HalcyonTM multileaf collimator system. *J Appl Clin Med Phys*. 2019;20:106-114.
27. Archambault Y, Boylan C, Bullock D, et al. Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning. *Med Phys Intl J*. 2020;8.
28. Wilcoxon F. Individual comparisons by ranking methods. *Breakthroughs in Statistics*. New York, NY: Springer; 1992:196-202.
29. Mohamed AS, Bahig H, Aristophanous M, et al. Prospective in silico study of the feasibility and dosimetric advantages of MRI-guided dose adaptation for human papillomavirus positive oropharyngeal cancer patients compared with standard IMRT. *Clin Transl Radiat Oncol*. 2018;11:11-18.
30. Harms J, Zhang J, Kayode O, et al. Implementation of a knowledge-based treatment planning model for cardiac-sparing lung radiation therapy. *Adv Radiat Oncol*. 2021;6: 100745.
31. Hansen CR, Crijns W, Hussein M, et al. Radiotherapy Treatment plannINg study Guidelines (RATING): A framework for setting up and reporting on scientific treatment planning studies. *Radiother Oncol*. 2020;153:67-78.