Published in final edited form as:

Radiother Oncol. 2016 August ; 120(2): 349–355. doi:10.1016/j.radonc.2016.06.010.

Evaluating Inter-Campus Plan Consistency Using a Knowledge Based Planning Model

Sean L Berry, Ph.D., Rongtao Ma, M.S., Amanda Boczkowski, M.S., Andrew Jackson, Ph.D., Pengpeng Zhang, Ph.D., and Margie Hunt, M.S.

Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Abstract

Background and Purpose—We investigate whether knowledge based planning (KBP) can identify systematic variations in intensity modulated radiotherapy (IMRT) plans between multiple campuses of a single institution.

Material and Methods—A KBP model was constructed from 58 prior main campus (MC) esophagus IMRT radiotherapy plans and then applied to 172 previous patient plans across MC and 4 regional sites (RS). The KBP model predicts DVH bands for each organ at risk which were compared to the previously planned DVH's for that patient.

Results—RS1's plans were the least similar to the model with less heart and stomach sparing, and more variation in liver dose, compared to MC. RS2 produced plans most similar to those expected from the model. RS3 plans displayed more variability from the model prediction but overall, the DVH's were no worse than those of MC. RS4 did not present any statistically significant results due to the small sample size (n=11).

Conclusions—KBP can retrospectively highlight subtle differences in planning practices, even between campuses of the same institution. This information can be used to identify areas needing increased consistency in planning output and subsequently improve consistency and quality of care.

Keywords

Knowledge based planning; quality assurance; radiation therapy; IMRT

Corresponding author: Sean L Berry, Ph.D., Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, Ph: 212-639-3305, Fax: 212-717-3258, BerryS@MSKCC.org.

CONFLICT OF INTEREST STATEMENT

A license for RapidPlan[®] KBP software was provided by Varian Medical Systems (VMS) to the authors as a part of a software evaluation agreement between Memorial Sloan Kettering Cancer Center and VMS. Sean Berry, Margie Hunt, and Pengpeng Zhang hold research grants from VMS unrelated to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

I. INTRODUCTION

Radiotherapy treatment plan variability has been observed between institutions[1–4] as well as between planners within a given institution[5, 6]. One method of reducing such variability is to define a comprehensive set of planning criteria for disease specific target coverage and normal tissue sparing. However, these are population based and do not account for the variability in patients' target and normal tissue geometry. When the geometry is such that the population based constraints are easily attainable, substandard plans may be produced if a planner does not appreciate that further decreases in critical organ doses are possible. Conversely, when the geometry is challenging, planners may waste time reaching for physically unachievable OAR sparing while retaining target coverage. Knowledge based planning (KBP) is a technique that utilizes a model trained from a database of prior plans to accurately predict DVH's for the current patient[7]. KBP accounts for tumor and normal tissue geometry and has been shown to reduce plan variability[1, 5, 8, 9].

A growing number of medical centers consist of a main campus and geographically distinct regional campuses. While all campuses may share segmentation guidelines, planning procedures, and evaluation criteria, traditional barriers to communication such as physical distance between sites and diminished face to face contact between staff members can reduce the dissemination of information and best practices[10]. The purpose of this study was to investigate whether a commercial knowledge based planning (KBP) product could be used to retrospectively identify systematic variation in intensity modulated radiation therapy (IMRT) plans created at multiple campuses of a single institution.

II. MATERIALS AND METHODS

Model Building

After obtaining appropriate IRB approval, 246 patients treated to the thoracic esophagus or gastroesophageal (GE) junction at any of our five facilities (main campus (MC) or four regional sites (RS)) between 2009 and 2014 were identified. All patients were planned for a prescribed dose of 50.4 Gy, delivered over 28 fractions, with IMRT and 6 MV x-rays using an in-house planning system[11, 12]. A standardized beam arrangement and well defined plan evaluation criteria (Table 1) were consistently in use during the stated time period. All plans were constructed and reviewed by the planners and physicians assigned to the campus where the patient was treated; there was no cross-campus planning. The patient population exhibited significant variability in tumor size and length and organ-at-risk (OAR) exposure, making this site ideal for KBP approaches. All organs-at-risk (OARs) were segmented in their anatomical entirety following standard atlases[13].

The first step of this study was the creation of a knowledge-based model that reflected our institutional best practice for the planning of esophageal cancer. All planning CTs, DICOM-RT Plans, Structure Sets, and Dose (2mm resolution) objects were transferred to Eclipse® (version 13.5, Varian Medical Systems, Palo Alto, CA) for analysis with the RapidPlan® (RP) KBP package which incorporates algorithms presented in the literature[9, 14]. The model was generated from plans from MC since those patients formed a large cohort treated by physicians highly experienced in GI malignancies. Although the model was based on MC

patients, the methodology (discussed below) did not preclude finding systematically superior results from another campus.

Sixty-four candidate plans for model creation were randomly chosen from an initial set of 92 MC plans. Six were subsequently removed due to irregular anatomy (e.g. severe scoliosis) or because they had been designed deliberately outside of the standard criteria (Table 1) for clinical reasons. The remaining 58 plans were used to train the RP model, which consisted of performing regression analysis to define a relationship between the target and OAR geometry and the resulting shapes of the OAR DVH's. Data points causing over-fitting of the particular OAR model, as indicated by their reported Cook's Distance[15], were reviewed and removed. Potential geometric and dosimetric outliers, identified using metrics provided by the software (modified Z-score, studentized residual, areal difference of estimate), were analyzed and their effect upon the model's reported coefficient of determination and chi-square statistics guided decisions about their removal. The final model contained 55 lungs and hearts, 56 stomachs, and 58 livers. The model was then validated using another 10 MC patients with a similar mix of target lengths and locations. Model regression plots, PTV and OAR volumes, and a summary of the planned dose/volume (D/V) parameters in the model and evaluation plans, are available online in the supplemental materials.

Model Application

The model was applied to the remaining 172 patients: 28 from MC, 38 from RS1, 58 from RS2, 37 from RS3, and 11 from RS4. When a RapidPlan[®] model is applied to a specific patient, upper and lower bounds representing ± 1 standard deviation (SD) about the predicted DVH are provided but the predicted DVH itself is not. The dashed lines in Figure 1 show examples of these prediction bands. For the analyses performed in this study, bins of the original planned (clinical) DVH were considered "outside" or "inside" of the prediction if they lay outside or inside of the bounds, respectively.

All planned and predicted DVH bands were imported into Matlab (version 2014b, Mathworks, Natick, MA). In addition to evaluating each clinical plan's D/V parameters against their predictions, three methods for more generally comparing planned and predicted DVHs were implemented. First, DVH's were compared bin-by-bin by calculating the percentage of the planned DVH bins that were outside the prediction bands (P):

$$P = (N_u + N_L) / (N_{tot}) * 100\% \quad (1)$$

where N_{tot} , N_U , and N_L are the total number of bins, the number above the upper bound, and below the lower bound, respectively. The bin size was 0.1 Gy. The percentage of bins worse than the prediction relative to the total number of bins outside either bound (F) was also calculated:

$$F = (N_u) / (N_u + N_L) * 100\% \quad (2)$$

P and F describe the number of planned DVH bins outside the prediction band but do not quantify the magnitude of the discrepancy between planned and predicted values. Appenzoller et al.[14] state that a standard sum of squared residuals method is not appropriate in this context because useful information about negative versus positive residuals is lost, leading to their introduction of the concept of the restricted sum of residuals (RSR). We expand on the RSR and present the modified RSR (mRSR), accounting for plans either better (negative values) or worse (positive values) than a predicted band, as the second analysis method:

$$\text{mRSR}_{ij} = \sum_{D=0}^{\infty} (\varepsilon_{ij,\text{upper}}^{+} * \Delta D) + \sum_{D=0}^{\infty} (\varepsilon_{ij,\text{lower}}^{-} * \Delta D) \quad (3)$$

given,

$$\varepsilon_{ij,\text{upper}} = \text{DVH}_{ij,\text{planned}} - \text{DVH}_{ij,\text{UpperEst}} \quad (4)$$

$$\varepsilon_{ij,\text{lower}} = \text{DVH}_{ij,\text{planned}} - \text{DVH}_{ij,\text{LowerEst}} \quad (5)$$

and,

$$\varepsilon_{ij,\text{upper}}^{+} = \varepsilon_{ij,\text{upper}}, \text{ if } \varepsilon_{ij,\text{upper}} > 0, \text{ else} = 0 \quad (6)$$

$$\varepsilon_{ij,\text{lower}}^{-} = \varepsilon_{ij,\text{lower}}, \text{ if } \varepsilon_{ij,\text{lower}} < 0, \text{ else} = 0 \quad (7)$$

The ε_{ij} values represent the difference in percent volume between the predicted and planned DVH at each given DVH bin for the j_{th} OAR of the i_{th} dataset. The ΔD values represent the width of the DVH bin.

The third analysis method utilizes the generalized equivalent uniform dose (gEUD)[16]. For each DVH structure, a gEUD was calculated for the upper predicted bound, the lower predicted bound, and the planned curve using the values of the parameter ‘a’ listed in Table 2. Figure 1 provides examples of P, F, mRSR, and gEUD for different levels of agreement between predicted and planned DVH’s.

All results were categorized by campus. The null hypothesis was that for each campus, the metrics and metric variability would be similar to that at MC, representing the institutional best practice. For the bin-to-bin DVH comparisons and mRSR, a statistically significant difference in the Student’s t-test ($p < 0.05$) indicated that the RS differed from the model more than the MC. For the gEUD and evaluation at specific D/V constraint points, two by

two contingency tables using Fischer's exact test were constructed to compare the number of patients inside versus outside of, and better versus worse than, the predicted bands between each RS and MC.

III. RESULTS

As shown in Tables 1–3, RS1 produced plans that were least similar to the model. Compared to MC, the heart V30_{Gy} was more often worse (76.2% vs 38.5%, $p=0.038$) and a greater percentage of heart DVH bins was worse than the prediction (74.2% vs 52.1%, $p=0.026$). For the stomach, the mRSR was worse (0.9 Gy vs –0.2 Gy, $p=0.029$) and the mean dose was more often outside the prediction (52.6% vs 14.3%, $p=0.002$), although there was no significant difference in the gEUD. There was also more deviation from the model, compared to MC, for the liver. The percentage of DVH bins outside the estimated bounds was greater (65.3% vs 49.7%, $p=0.004$) and the mean dose was outside of the predicted bounds more often (60.5% vs 32.1%, $p=0.027$).

RS2 produced plans most similar to those expected from the model. Aside from the number of patients outside the lungs V10_{Gy} prediction (53.4% vs 25.0%, $p=0.020$), the stomach was the only organ with a statistically significant difference. The stomach mean dose, gEUD, and mRSR were worse than the predictions more often than MC (82.8% vs 25.0%, $p=0.036$, 86.7% vs 37.5%, $p=0.026$, and 0.7 Gy vs –0.2 Gy, $p=0.002$, respectively).

RS3 plans varied from the model predictions more than those from MC. For the heart and stomach, the percentage of DVH bins outside the estimated bounds (57.3% vs 42.1%, $p=0.009$ and 54.6% vs 31.8%, $p<0.001$, respectively), percentage of plans outside of the gEUD bounds (59.5% vs 21.4%, $p=0.003$ and 59.5% vs 28.6%, $p=0.012$, respectively), and percentage of plans with mean doses outside the bounds (56.8% vs 21.4%, $p=0.005$ and 56.8% vs 14.3%, $p=0.001$, respectively) was greater. For the lungs and liver, the percentage of DVH bins outside of the estimated bounds (56.3% vs 40.5%, $p=0.004$ and 63.8% vs 49.7%, $p=0.013$, respectively) was greater but there was no difference in terms of gEUD or mRSR. Further, for the lungs V40_{Gy} and V10_{Gy} more plans were outside the predicted D/V parameters (59.5% vs 32.1%, $p=0.045$ and 62.2% vs 25.0%, $p=0.005$, respectively), although they were not significantly better or worse.

Due to the small sample size ($n=11$), no statistically significant differences from MC were apparent for RS4.

IV. DISCUSSION

KBP has been commercially implemented as an automated optimization tool[8, 18] but has also shown to be an effective tool for comparing treatment plans against historical controls[6, 14, 17]. Moore et al.[19] has found, using NTCP analysis, that KBP can identify suboptimal plans in large interinstitutional trials even when the trials included rigorous segmentation guidelines and dosimetric OAR constraints. Such variability has been attributed to differences in planning systems, linac capabilities, and institutional experience with the treatment technique[20, 21] but not, to date, individual planner characteristics such as education, certification status, or years of experience[10, 22]. In this study, we were

interested in determining whether KBP could identify systematic differences in planning approaches between campuses of a single institution, where again, the same planning system, planning guidelines, and plan evaluation criteria were used. Our hypothesis, similar to that of Moore et al, was that a KBP approach would be more sensitive to subtle plan differences than an approach based solely on compliance with planning criteria. A KBP approach could implicitly account for patient anatomy differences and could evaluate the entire DVH shape. Used in this manner, KBP could help identify sources of planning inconsistency across large facilities; for example, differences in physician OARsparing preferences or weakness in staff training or communication. Our focus, at this stage, was solely to determine if KBP could be used to identify plan differences and not to identify their cause. KBP did detect systematic differences in the heart gEUD, liver mean dose, and stomach mRSR, indicating that subtle differences in the planning approach at our various campuses do exist. Furthermore, the MC cohort most closely resembled the model, most likely reflecting a higher level of consistency amongst physicians specializing in GI malignancies and working closely together at the MC that is difficult to reproduce in the more disperse practices at the RS. Further investigation will include more in-depth evaluation of these and other issues. However, it has already been determined that a larger fraction (3/5) of planners at RS2, whose plans were closest to the model, were originally employed at MC than those at RS1 (1/3) or RS3 (2/5). Furthermore, physicians at RS3 often request that the heart dose be kept to the higher end of the guideline range (mean <30 Gy), foregoing the V30_{Gy}, while keeping the lungs V10_{Gy} to the lower end (closer to 35%), even though the order of priorities in Table 1 lists the opposite. These observations demonstrate the importance of training, communication, and continuing collaboration between all clinical staff and have led to changes in our cross-campus planner educational programs.

Despite some differences between facilities, this analysis highlighted an overall consistency with our institutional planning priorities for this anatomical site. Table 1 shows that meeting the V20_{Gy}, V30_{Gy}, and V40_{Gy} lung dose constraints is the top clinical priority for organs at risk at our institution. Commensurate with that, the KBP results from all campuses showed that lungs doses were the least variable. Conversely, sparing the stomach is a lower planning priority and we observed that P, mRSR, and gEUD were worse for 3, 2, and 2 of the RS's respectively, indicating a need for better defined and communicated planning criteria for the stomach. Such data provides valuable feedback to refine our clinical plan evaluation criteria (Table 1), as well as educational and communication processes to improve consistency.

Thoroughly comparing DVH's and dose distributions beyond the analysis of a few D/V points is challenging. We used both DVH-shape and biological based metrics, similar to approaches utilized by others[14, 19], and found these metrics to be helpful in identifying DVH's that differed considerably from the predicted bounds. Figure 2 presents DVHs for a RS2 patient with heart, liver, and stomach metrics that were noticeably high (Table 4). Re-planning the patient, incorporating the predictions from KBP, successfully lowered these OAR doses, and corresponding DVH metrics, while maintaining good PTV coverage.

There are limitations to this analysis and the use of KBP for the evaluation of plan consistency. Most importantly, the DVH variability built into the model must be considered and closely evaluated during the model building process[23]. For this project, we wished to

keep the prediction bands fairly narrow in order to reflect the institutional best practice as closely as possible. Secondly, as this is a retrospective study, the patient cohorts from each facility could not be exactly matched in terms of anatomical or geometrical variability (Supplemental Table 1). Both of these factors could affect the sensitivity and specificity of the KBP approach for consistency evaluation. Finally, due to limitations in the commercial software, our analysis was based solely on the use of DVH prediction bands ± 1 standard deviation about the DVH estimate, rather than the estimate itself. Use of the latter would have allowed more flexibility in choice of metrics and subsequent analysis.

V. CONCLUSIONS

Knowledge based treatment planning can be used to analyze the consistency of treatment plans between campuses of a single institution. Our analysis naturally highlighted and demonstrated planning priorities for esophageal IMRT; lung sparing as most critical, stomach sparing as less critical. It also allowed us to identify educational, communication and procedural areas that we can focus on to increase consistency. The methods used in this study could be generalized and repeated for any treatment site of interest.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to thank the planning staffs of the main campus and regional centers, and their respective chiefs, for their cooperation and interest in this study. This research was partially supported by the MSK Cancer Center Support Grant/Core Grant (P30 CA008748).

References

1. Good D, Lo J, Lee WR, Wu QJ, Yin FF, Das SK. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. *International journal of radiation oncology, biology, physics*. 2013; 87:176–81.
2. Das IJ, Cheng CW, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems. *J Natl Cancer I*. 2008; 100:300–7.
3. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: An interinstitutional study of planners and planning systems. *Practical radiation oncology*. 2012; 2:296–305. [PubMed: 24674168]
4. Lian J, Yuan L, Ge Y, et al. Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: an intertechnique and interinstitutional study. *Medical physics*. 2013; 40:121704. [PubMed: 24320490]
5. Moore KL, Brame RS, Low DA, Mutic S. Experience-based quality control of clinical intensitymodulated radiotherapy planning. *International journal of radiation oncology, biology, physics*. 2011; 81:545–51.
6. Wu B, Ricchetti F, Sanguineti G, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Medical physics*. 2009; 36:5497–505. [PubMed: 20095262]
7. Tol JP, Dahele M, Delaney AR, Slotman BJ, Verbakel WF. Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? *Radiation Oncology*. 2015; 10:1. [PubMed: 25567003]

8. Wu B, Ricchetti F, Sanguineti G, et al. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. *International journal of radiation oncology, biology, physics*. 2011; 79:1241–7.
9. Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Medical physics*. 2012; 39:6868–78. [PubMed: 23127079]
10. Berry SL, Boczkowski A, Ma R, Mechalakos J, Hunt M. Interobserver Variability in Radiotherapy Plan Output: Results of a Single-Institution Study. *Practical radiation oncology*.
11. Mohan R, Barest G, Brewster LJ, et al. A comprehensive three-dimensional radiation treatment planning system. *International Journal of Radiation Oncology* Biology* Physics*. 1988; 15:481–95.
12. Spirou SV, Chui C-S. A gradient inverse planning algorithm with dose-volume constraints. *Medical physics*. 1998; 25:321–33. [PubMed: 9547499]
13. Jabbour SK, Hashem SA, Bosch W, et al. RTOG Consensus Panel: Upper Abdominal Normal Organ Contouring Guidelines. 2013
14. Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Medical physics*. 2012; 39:7446–61. [PubMed: 23231294]
15. Cook RD. Detection of influential observation in linear regression. *Technometrics*. 1977:15–8.
16. Niemierko A. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Medical physics*. 1997; 24:103–10. [PubMed: 9029544]
17. Zhu X, Ge Y, Li T, Thongphiew D, Yin FF, Wu QJ. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Medical physics*. 2011; 38:719–26. [PubMed: 21452709]
18. Petit SF, Wu BB, Kazhdan M, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. *Radiother Oncol*. 2012; 102:38–44. [PubMed: 21680036]
19. Moore KL, Schmidt R, Moiseenko V, et al. Quantifying Unnecessary Normal Tissue Complication Risks due to Suboptimal Planning: A Secondary Study of RTOG 0126. *International journal of radiation oncology, biology, physics*. 2015; 92:228–35.
20. Bohsung J, Gillis S, Arrans R, et al. IMRT treatment planning - A comparative inter-system and intercentre planning exercise of the QUASIMODO group. *Radiotherapy and Oncology*. 2005; 76:354–61. [PubMed: 16154218]
21. Williams MJM, Bailey MJ, Forstner D, Metcalfe PE. Multicentre quality assurance of intensitymodulated radiation therapy plans: A precursor to clinical trials. *Australasian Radiology*. 2007; 51:472–9. [PubMed: 17803801]
22. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Practical Radiation Oncology*. 2012; 2:296–305. [PubMed: 24674168]
23. Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WF. Evaluation of a knowledge-based planning solution for head and neck cancer. *International journal of radiation oncology, biology, physics*. 2015; 91:612–20.
24. Martel MK, Sahijdak WM, Ten Haken RK, Kessler ML, Turrisi AT. Fraction size and dose parameters related to the incidence of pericardial effusions. *International journal of radiation oncology, biology, physics*. 1998; 40:155–61.
25. Burman C, Kutcher GJ, Emami B, Goitein M. Fitting of normal tissue tolerance data to an analytic function. *International journal of radiation oncology, biology, physics*. 1991; 21:123–35.
26. Dawson LA, Normolle D, Balter JM, McGinn CJ, Lawrence TS, Ten Haken RK. Analysis of radiationinduced liver disease using the Lyman NTCP model. *International journal of radiation oncology, biology, physics*. 2002; 53:810–21.

Figure 1.

Typical values of P (percentage of bins of the planned DVH that exceeded the predicted bounds), F (percentage of bins exceeding the upper bound relative to the total number of bins exceeding either bound), mRSR (modified restricted sum of residuals), and gEUD (shown in all panels of this example: heart, $a = 1.6$) for different relationships between the planned DVH (solid line) and predicted DVH bounds (dashed lines).

Figure 2.

The clinical planned DVH (solid black line), predicted DVH bounds (dashed lines), and replanned DVH (solid gray line) for patient 291 at RS2 where the DVH metrics were noticeably high. Re-planning resulted in DVH's much closer to their estimates.

Table 1

Summary of clinical constraints used to generate and evaluate esophageal IMRT treatment plans, listed in order from highest to lowest planning priority in column 1. Columns 3–8 are a tabulation of the number of clinical plans “inside” versus “outside” and “better” versus “worse” than their corresponding predicted DVH bands at each of the clinical dose/volume evaluation points. Bolded and italicized values represent regional sites that statistically differed ($p < 0.05$) from MC for that metric.

Location of Planned DVH Relative to Predicted Bounds											Quality of Planned DVH Relative to Prediction			Better vs Worse
Organ / Criteria		Campus	# (%) inside	# (%) outside	Inside Vs. Outside		# (%) better		# (%) worse	p-value				
Lungs V20 _{Gy}	20%	MC	14 (50.0%)	14 (50.0%)			6 (42.9%)	8 (57.1%)						
		RS1	25 (65.8%)	13 (34.2%)	0.217		8 (61.5%)	5 (38.5%)			0.450			
		RS2	40 (69.0%)	18 (31.0%)	0.101		10 (55.6%)	8 (44.4%)			0.722			
		RS3	16 (43.2%)	21 (56.8%)	0.311		5 (23.8%)	16 (76.2%)			0.283			
		RS4	7 (63.6%)	4 (36.4%)	0.497		2 (50.0%)	2 (50.0%)			1.000			
Lungs V30 _{Gy}	15%	MC	17 (60.7%)	11 (39.3%)			7 (63.6%)	4 (36.4%)						
		RS1	20 (52.6%)	18 (47.4%)	0.618		11 (61.1%)	7 (38.9%)			1.000			
		RS2	31 (53.4%)	27 (46.6%)	0.644		19 (70.4%)	8 (29.6%)			0.714			
		RS3	16 (43.2%)	21 (56.8%)	0.213		12 (57.1%)	9 (42.9%)			1.000			
		RS4	6 (54.5%)	5 (45.5%)	0.734		3 (60.0%)	2 (40.0%)			1.000			
Lungs V40 _{Gy}	10%	MC	19 (67.9%)	9 (32.1%)			7 (77.8%)	2 (22.2%)						
		RS1	20 (52.6%)	18 (47.4%)	0.311		10 (55.6%)	8 (44.4%)			0.406			
		RS2	26 (44.8%)	32 (55.2%)	0.065		24 (75.0%)	8 (25.0%)			1.000			
		RS3	15 (40.5%)	22 (59.5%)	0.045		15 (68.2%)	7 (31.8%)			0.689			
		RS4	5 (45.5%)	6 (54.5%)	0.277		4 (66.7%)	2 (33.3%)			1.000			
Heart V30 _{Gy}	20–30%	MC	15 (53.6%)	13 (46.4%)			8 (61.5%)	5 (38.5%)						
		RS1	17 (44.7%)	21 (55.3%)	0.619		5 (23.8%)	16 (76.2%)			0.038			
		RS2	32 (55.2%)	26 (44.8%)	1.000		13 (50.0%)	13 (50.0%)			0.734			
		RS3	13 (35.1%)	24 (64.9%)	0.206		13 (54.2%)	11 (45.8%)			0.739			
		RS4	2 (18.2%)	9 (81.8%)	0.073		4 (44.4%)	5 (55.6%)			0.666			
Heart Mean	30 Gy	MC	22 (78.6%)	6 (21.4%)			1 (16.7%)	5 (83.3%)						
		RS1	13 (34.2%)	25 (65.8%)	0.001		6 (24.0%)	19 (76.0%)			1.000			

Location of Planned DVH Relative to Predicted Bounds										Quality of Planned DVH Relative to Prediction		Better vs Worse
Organ / Criteria	Campus	# (%) inside	# (%) outside	p-value	# (%) better	# (%) worse	p-value					
Liver V20 _{Gy} 20–45%	RS2	39 (67.2%)	19 (32.8%)	0.321	8 (42.1%)	11 (57.9%)	0.364					
	RS3	16 (43.2%)	21 (56.8%)	0.005	9 (42.9%)	12 (57.1%)	0.363					
	RS4	7 (63.6%)	4 (36.4%)	0.424	2 (50.0%)	2 (50.0%)	0.500					
	MC	10 (35.7%)	18 (64.3%)		11 (61.1%)	7 (38.9%)						
Liver V20 _{Gy} 20–45%	RS1	8 (21.1%)	30 (78.9%)	0.264	15 (50.0%)	15 (50.0%)	0.555					
	RS2	21 (36.2%)	37 (63.8%)	1.000	17 (45.9%)	20 (54.1%)	0.391					
	RS3	13 (35.1%)	24 (64.9%)	1.000	12 (50.0%)	12 (50.0%)	0.542					
	RS4	2 (18.2%)	9 (81.8%)	0.446	9 (100.0%)	0 (0.0%)	0.059					
Liver V30 _{Gy} 20%	MC	13 (46.4%)	15 (53.6%)		9 (60.0%)	6 (40.0%)						
	RS1	12 (31.6%)	26 (68.4%)	0.305	13 (50.0%)	13 (50.0%)	0.746					
	RS2	25 (43.1%)	33 (56.9%)	0.819	19 (57.6%)	14 (42.4%)	1.000					
	RS3	8 (21.6%)	29 (78.4%)	0.060	21 (72.4%)	8 (27.6%)	0.501					
Liver V30 _{Gy} 20%	RS4	1 (9.1%)	10 (90.9%)	0.060	8 (80.0%)	2 (20.0%)	0.402					
	MC	19 (67.9%)	9 (32.1%)		4 (44.4%)	5 (55.6%)						
	RS1	15 (39.5%)	23 (60.5%)	0.027	12 (52.2%)	11 (47.8%)	1.000					
	RS2	34 (56.6%)	24 (41.4%)	0.482	10 (41.7%)	14 (58.3%)	1.000					
Liver V30 _{Gy} 20%	RS3	25 (67.6%)	12 (32.4%)	1.000	6 (50.0%)	6 (50.0%)	1.000					
	RS4	6 (54.5%)	5 (45.5%)	0.478	4 (80.0%)	1 (20.0%)	0.301					
	MC	24 (85.7%)	4 (14.3%)		3 (75.0%)	1 (25.0%)						
	RS1	18 (47.4%)	20 (52.6%)	0.002	6 (30.0%)	14 (70.0%)	0.130					
Stomach Mean 30 Gy	RS2	29 (50.0%)	29 (50.0%)	0.005	5 (17.2%)	24 (82.8%)	0.036					
	RS3	16 (43.2%)	21 (56.8%)	0.001	9 (42.9%)	12 (57.1%)	0.322					
	RS4	9 (81.8%)	2 (18.2%)	1.000	1 (50.0%)	1 (50.0%)	1.000					
	MC	21 (75.0%)	7 (25.0%)		6 (85.7%)	1 (14.3%)						
Lungs V10 _{Gy} 35–65%	RS1	18 (47.4%)	20 (52.6%)	0.042	10 (50.0%)	10 (50.0%)	0.183					
	RS2	27 (46.6%)	31 (53.4%)	0.020	15 (48.4%)	16 (51.6%)	0.104					
	RS3	14 (37.8%)	23 (62.2%)	0.005	17 (73.9%)	6 (26.1%)	1.000					
	MC	21 (75.0%)	7 (25.0%)		6 (85.7%)	1 (14.3%)						

Organ / Criteria	Location of Planned DVH Relative to Predicted Bounds			Inside Vs. Outside		Quality of Planned DVH Relative to Prediction		Better vs Worse	
	Campus	# (%) inside	# (%) outside	p-value		# (%) better	# (%) worse	p-value	
	RS4	6 (54.5%)	5 (45.5%)	0.262		4 (80.0%)	1 (20.0%)	1.000	

Table 2

Tabulation of the number of plans “inside” versus “outside” and “better” versus “worse” than the gEUD predicted values for each organ at each campus. The following values of the parameter ‘a’ were chosen for the gEUD calculations: heart (1.6)[24], lungs (1.0)[25], and stomach (6.7)[25]. The liver is not included in this table because an a = 1.0 was recommended[26], but this corresponds to the mean dose shown in Table 1. Bolded and italicized values represent regional sites that statistically differed ($p < 0.05$) from MC for that metric. A patient is scored as “inside” if the planned gEUD was between the gEUDs of the upper and lower predicted bounds. For those that were outside the bounds, they were recorded as “better” if the planned gEUD was less than the lower predicted gEUD bound or “worse” if it was higher than the upper predicted gEUD bound.

Organ	Campus	gEUD Predicted bounds		Inside vs Outside	gEUD Predicted bounds		Better vs Worse
		# (%) inside	# (%) outside	p-value	# (%) better	# (%) worse	p-value
Lungs (a = 1.0)							
	MC	24 (85.7%)	4 (14.3%)		2 (50.0%)	2 (50.0%)	
	RS1	29 (76.3%)	9 (23.7%)	0.533	6 (66.7%)	3 (33.3%)	1.000
	RS2	41 (70.7%)	17 (29.3%)	0.182	8 (47.1%)	9 (52.9%)	1.000
	RS3	24 (64.9%)	13 (35.1%)	0.087	9 (69.2%)	4 (30.8%)	0.584
	RS4	8 (72.7%)	3 (27.3%)	0.379	2 (66.7%)	1 (33.3%)	1.000
Heart (a = 1.6)							
	MC	22 (78.6%)	6 (21.4%)		1 (16.7%)	5 (83.3%)	
	RS1	15 (39.5%)	23 (60.5%)	0.002	6 (26.1%)	17 (73.9%)	1.000
	RS2	39 (67.2%)	19 (32.8%)	0.321	8 (42.1%)	11 (57.9%)	0.364
	RS3	15 (40.5%)	22 (59.5%)	0.003	10 (45.5%)	12 (54.5%)	0.355
	RS4	5 (45.5%)	6 (54.5%)	0.062	3 (50.0%)	3 (50.0%)	0.546
Stomach (a = 6.7)							
	MC	20 (71.4%)	8 (28.6%)		5 (62.5%)	3 (37.5%)	
	RS1	20 (52.6%)	18 (47.4%)	0.137	5 (27.8%)	13 (72.2%)	0.189
	RS2	43 (74.1%)	15 (25.9%)	0.800	2 (13.3%)	13 (86.7%)	0.026
	RS3	15 (40.5%)	22 (59.5%)	0.012	11 (50.0%)	11 (50.0%)	0.689
	RS4	7 (63.6%)	4 (36.4%)	0.708	1 (25.0%)	3 (75.0%)	0.546

Table 3

Average and standard deviation results for the DVH shape-based metrics for each organ at each campus. Bolded and italicized values represent regional sites that statistically differed ($p < 0.05$) from MC for that metric. The lungs displayed the most consistency and the stomach the least consistency between each RS and MC.

Organ	Campus	P = % DVH bins outside predicted band (avg +/- SD)(%) [p-value]	F = % DVH bins outside worse than prediction (avg +/- SD) (%) [p-value]	mRSR = net DVH area between planned and predicted DVH (avg +/- SD) (Gy) [p-value]
Lungs				
	MC	40.5 +/- 21.0	60.8 +/- 33.7	0.0 +/- 0.5
	RS1	49.0 +/- 21.0 [0.108]	57.4 +/- 34.9 [0.694]	-0.1 +/- 0.8 [0.328]
	RS2	47.8 +/- 19.2 [0.127]	53.6 +/- 38.3 [0.382]	0.0 +/- 0.5 [0.765]
	RS3	56.3 +/- 21.0 [0.004]	48.9 +/- 34.3 [0.169]	-0.2 +/- 0.7 [0.217]
	RS4	51.8 +/- 26.8 [0.228]	41.7 +/- 32.9 [0.123]	-0.2 +/- 0.5 [0.320]
Heart				
	MC	42.1 +/- 22.6	52.1 +/- 40.6	0.3 +/- 1.5
	RS1	54.9 +/- 27.8 [0.044]	74.2 +/- 36.2 [0.026]	0.7 +/- 1.8 [0.292]
	RS2	48.6 +/- 22.1 [0.215]	58.2 +/- 35.7 [0.505]	0.0 +/- 1.8 [0.430]
	RS3	57.3 +/- 21.6 [0.009]	50.4 +/- 35.9 [0.960]	0.9 +/- 3.7 [0.323]
	RS4	47.0 +/- 19.2 [0.502]	57.8 +/- 37.3 [0.681]	0.1 +/- 1.4 [0.681]
Liver				
	MC	49.7 +/- 21.3	61.1 +/- 37.9	0.0 +/- 0.6
	RS1	65.3 +/- 19.3 [0.004]	63.2 +/- 34.7 [0.815]	0.0 +/- 1.5 [0.958]
	RS2	54.9 +/- 20.8 [0.290]	63.3 +/- 36.2 [0.797]	0.2 +/- 1.0 [0.476]
	RS3	63.8 +/- 22.5 [0.013]	56.5 +/- 32.9 [0.609]	0.2 +/- 0.8 [0.507]
	RS4	63.8 +/- 18.7 [0.055]	36.4 +/- 36.7 [0.076]	-0.9 +/- 1.6 [0.099]
Stomach				
	MC	31.8 +/- 22.0	55.3 +/- 43.5	-0.2 +/- 0.9
	RS1	49.8 +/- 23.4 [0.002]	72.2 +/- 40.7 [0.115]	0.9 +/- 2.8 [0.029]
	RS2	46.7 +/- 23.4 [0.005]	73.8 +/- 35.9 [0.058]	0.7 +/- 1.5 [0.002]
	RS3	54.6 +/- 21.1 [0.001]	61.7 +/- 42.5 [0.559]	0.4 +/- 2.7 [0.226]
	RS4	41.7 +/- 23.3 [0.241]	68.4 +/- 41.5 [0.394]	0.1 +/- 0.5 [0.232]

DVH metrics for the clinical plan and the Rapidplan[®] re-plan for patient 291 at RS2. While the metrics were noticeably high for the original plan, re-planning resulted in values much closer to the model prediction.

Table 4

Patient #291		Clinical	Replan
PTV	Max (Gy)	56.9	53.2
	D95% (Gy)	50.7	50.4
Lungs	V40Gy (%)	2.2	1.4
	V30Gy (%)	6.1	3.1
	V20Gy (%)	16.2	11.8
	V10Gy (%)	24.7	19.1
	P (%)	54.3%	68.1%
	F (%)	66.4%	8.7%
	gEUD (Gy)	8.1	7.3
	mRSR (Gy)	0.3	-0.3
Heart	V30Gy (%)	28.2	17.7
	Mean (Gy)	23.1	18.7
	P (%)	98.3%	23.5%
	F (%)	100%	56.5%
	gEUD (Gy)	24.1	20.7
	mRSR (Gy)	2.3	0.0
Liver	V30Gy (%)	11.4	6.3
	V20Gy (%)	40.7	33.3
	Mean (Gy)	14.8	12.5
	P (%)	84.2%	27.4%
	F (%)	100%	100%
	mRSR (Gy)	2.1	0.5
Stomach	Mean (Gy)	9.1	7.4
	P (%)	65.9%	19.3%
	F (%)	100.0%	51.0%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Patient #291	Clinical	Replan
gEUD (Gy)	43.6	41.4
mRSR (Gy)	2.5	0.1