Multi-institutional Validation of a Knowledge-based Planning Model for Patients Enrolled on RTOG 0617: Implications for Plan Quality Controls in Cooperative Group Trials

James A. Kavanaugh, MS[1], **Sarah Holler**[2], **Todd A DeWees, PhD**[1], **Clifford G. Robinson, MD**[1], **Jeffrey D. Bradley, MD**[1], **Puneeth Iyengar, MD, PhD**[3], **Kristin A. Higgins, MD**[4], **Sasa Mutic, PhD**[1], **Lindsey A. Olsen, PhD**[1]

[1]Department of Radiation Oncology, Washington University in St. Louis, St. Louis, MO, USA

[2]Department of Physics, Bucknell University, Lewisburg, PA, USA

[3]Department of Radiation Oncology, University of Texas Southwestern, Dallas, TX, USA

[4]Department of Radiation Oncology, Emory University, Atlanta, GA, USA

Abstract

Purpose—To evaluate the feasibility of using a single-institution knowledge based planning (KBP) model as a dosimetric plan quality control (QC) for multi-institutional clinical trials. The efficacy of this QC tool was retrospectively evaluated using a subset of plans submitted to RTOG 0617.

Methods and Materials—A single KBP model was created utilizing a commercially available software (RapidPlan™, Varian Medical Systems, Palo Alto, CA) and data from 106 patients with non-small cell lung cancer (NSCLC) treated at a single institution. All plans had prescriptions ranging from 60Gy/30fx to 74Gy/37fx and followed planning guidelines from RTOG 0617. Two sets of optimization objectives were created to produce different trade-offs using the single KBP model predictions: one prioritizing target coverage (PC) and a second prioritizing lung sparing (LS) while allowing acceptable variation in target coverage. Three institutions that submitted a high volume of clinical plans to RTOG 0617 provided 25 patients which were replanned using both sets of optimization objectives. Model-generated dose volume histogram predictions were used to identify patients that exceeded Lungs-CTV $V_{20Gy} > 37\%$ and would benefit from the LS objectives. Overall plan quality differences between KBP-generated plans and clinical plans were evaluated at RTOG 0617 defined dosimetric end-points.

Results—Target coverage and OAR sparing was significantly improved for most KBP generated plans compared to clinical trial data. The KBP model using PC objectives reduced heart D_{mean} and

Author to whom correspondence should be addressed: James A. Kavanaugh, MS, Department of Radiation Oncology, Washington University in Saint Louis, 4921 Parkview Place, Saint Louis, MO 63108, Tel: 314-252-2428, jkavanaugh@wustl.edu. **Author to whom statistical analysis questions should be addressed:** Todd A. DeWees, Department of Radiation Oncology, Washington University in Saint Louis, 4511 Forest Park Avenue, Saint Louis, MO 63108, Tel: 314-747-0059, tdewees@wustl.edu.

V$_{40Gy}$ by 2.1Gy and 5.2%, respectively. Similarly, utilizing LS objectives reduced the Lungs-CTV D$_{mean}$ and V$_{20Gy}$ by 2.0Gy and 2.9% respectively. The KBP predictions correctly identified all patients with Lungs-CTV V20Gy>37% (5/25) and significantly reduced dose to Lungs-CTV by applying LS optimization objectives.

Conclusions—A single institution KBP model can be applied as a QC tool for multi-institutional clinical trials to improve overall plan quality and provide decision-support to determine the need for anatomy-based dosimetric trade-offs.

Introduction

Knowledge-based planning (KBP) has become a prominent area of research in Radiation Oncology in the last decade. The development of KBP aims to address the lack of systematic quality control and plan quality variability in radiotherapy treatment planning [1–2]. Several investigators developed independent KBP methods with the common theme of using the knowledge of patient geometry and prior radiotherapy treatment data to improve the efficiency, standardization, and quality of the treatment planning process[3–6]. Recent approaches utilized model-based methods which derive correlations between patient geometry and delivered dose to predict achievable DVHs used as patient-specific IMRT optimization objectives[3,6]. The implementation of a KBP method into a commercially available treatment planning system allows for the use of KBP models in clinical practice and many studies have focused on training site-specific clinical models to predict achievable DVHs for individual patients [7–10].

The impact of KBP models to standardize and improve plan quality becomes evident when evaluating IMRT treatment plans created across multiple institutions. A study performed by Fogliata et al. demonstrated a systematic reduction in clinically relevant mean dose to the heart and lungs when an esophageal KBP model trained using plans from two institutions was applied to clinically treated plans at a third institution [11]. This inter-institutional variability in plan quality has been shown to exist even when strict planning guidelines are provided, as evidenced by a secondary analysis of RTOG 0126 performed by Moore et al. that correlated sub-optimal plan quality with excess risk of rectal toxicity[12]. In their study, Moore et al. demonstrated that global KBP models developed by cross-institutional groups using large-scale clinical trial data may potentially standardize the treatment plan quality for plans submitted to clinical trials, possibly resulting in improved patient outcomes.

One clinical trial that may have exhibited variable treatment plan quality impacting overall survival was RTOG 0617[13]. RTOG 0617 compared overall survival of patients with inoperable stage III NSCLC enrolled on standard-dose (60 Gy in 30 fractions) and high-dose (74 Gy in 37 fractions) arms, concluding that high-dose radiation was not beneficial and may be potentially harmful[14]. The primary analysis hypothesized a combination of inferior target coverage and increased dose to the heart in the high-dose arm may have contributed to the lower overall survival. The possibility of sub-optimal plan quality impacting patient outcomes was further investigated in a recent secondary analysis which showed treatments at institutions with lower clinical trial accrual was associated with reduced overall survival, with one hypothesis that the limited planning experience lead to a higher incidence of sub-

optimal treatment plans[13]. Based on the reasons outlined above, it has been hypothesized that utilizing a KBP model as a QC tool may improve quality and consistency of treatment planning for patients enrolled on RTOG 0617.

In this study we performed a multi-institutional validation of a stage III lung cancer KBP model trained with patients treated at a single institution (Institution A) with patients enrolled on RTOG 0617 treating at three participating institutions. The first objective was to show that a KBP model trained with patients enrolled on an investigator initiated trial at a single institution can be used as a QC to improve dosimetric plan quality for patients enrolled on a multi-institutional clinical trial. Additionally, this study investigated the efficacy of using the KBP model generated DVH predictions to evaluate whether a treatment plan can achieve the defined clinical trial goals.

Methods and Materials

Initial Model Training

The initial KBP lung model was trained using a patient dataset consisting of 106 stage III NSCLC clinical plans treated at institution A that satisfied the contouring guidelines of RTOG 0617. All plans in the training dataset represent manually created, clinically treated patients from 2009–2013 delivered with 6 MV static field IMRT using six-to-eight beam angles selected based on each individual patient's anatomy. Planning optimization objectives for all patients followed institutional planning guidelines which differed slightly from those provided by RTOG 0617, with RTOG 0617 and institutional goals given in Table 1. All plans were calculated in the Eclipse (v13.6.23) treatment planning system using the photon optimizer (PO) and anisotropic analytical algorithm (AAA) with a 2.5mm dose grid. The primary difference between institutional and RTOG 0617 planning goals is in PTV coverage goals, with institution A allowing an acceptable variation of 95% of the target covered by 95% of the prescription, with a consequence of lower OAR predictions. PTV coverage in the KBP model achieved RTOG 0617 planning goals by setting model specific PTV optimization objectives to increase coverage, while still utilizing the lower OAR predictions to increase dosimetric sparing. The initial model was generated using primary targets located in left lung (31 patients), right lung (52 patients), and mediastinum/bilateral (23 patients). Figure 1 details the entire study methodology.

An extensive model refinement process was conducted to remove any dosimetrically sub-optimal treatment plans. Outliers were initially identified for each of the OARs outside one standard deviation on each OAR principle component regression plot and that exceeded at least one of the 4 RapidPlan provided statistical evaluation metrics, with thresholds given in parenthesis: Cook's Distance (10), Studentized Residual (3), Modified Z-score (3.5), and Areal Difference of Estimate (3). A total of 19 patients were identified as having 2 or more OAR dosimetric outliers and all 19 were manually replanned to improve plan quality. In the replanning process, the identified outlier OARs were prioritized to achieve maximize dosimetric plan improvement, while maintaining at least 95% PTV coverage by 95% of the prescription and achieved dosimetric sparing to the non-outlier OARs. All 19 plans were combined with the original remaining 87 patient plans, data was re-extracted, and a refined model was trained. A minimal number of individual OARs still above 1 standard deviation

from the respective regression plots were removed from the training set and a final lung model (FLM) was retrained with the remaining data. A qualitative comparison of predictions produced by the initial lung model and FLM for the single institution validation dataset indicated reduction in OAR DVH predictions and narrower bands of prediction uncertainty.

Single Institution Model Validation

An independent patient validation set of 9 cases evenly divided between left lung, right lung, and bilateral/mediastinum was identified from the pool of clinically delivered plans treated at Institution A. During validation, optimization objectives prioritizations for each target and OAR DVH metric were iteratively and systematically refined for the entire validation set to ensure target coverage goals were precisely achieved through optimization for the validation population while maximizing OAR sparing as specified in the clinical planning priorities in RTOG 0617. Two unique sets of optimization objectives were created: PTV Coverage (PC) and Lung Sparing (LS) (Table 1). The PC optimization objectives were weighted to produce treatment plans with 100% of the prescription covering exactly 95% of the target while maximizing OAR sparing. The LS optimization objectives were modified to generate treatment plans that achieved acceptable variation in target coverage (prescription covering exactly 90% of the target), while further reducing the Lungs-CTV mean dose and V_{20Gy}. Iterative refinement of the optimization objective priorities was completed when the average target coverage by the prescription matched PC and LS goals within 0.5%, with less than 1% plan normalization required for any individual plan to achieve the respective target coverages. Additionally all OAR DVHs produced from the model were required to fall between midline and the lower bound of the prediction range for the RTOG 0617 clinically defined DVH end-points. Once appropriately balanced optimization objectives priorities were determined, they were set within the model and applied to all cases without manual intervention.

Multi-Institutional RTOG 0617 Model Validation

A dataset of 25 patients submitted to RTOG 0617 from three institutions were used to validate the FLM, with Institution A submitting 9 plans (6 right-sided, 3 left-sided, 1 mediastinal), Institution B submitting 10 plans (5 right-sided, 2 left-sided, 3 mediastinal), and Institution C submitting 6 plans (1 left-sided, 5 mediastinal). All multi-institutional datasets were independent of the single institutional validation datasets. The majority of plans (22/25) were delivered using static field IMRT, with the remaining 3 plans delivered with volumetric modulated arc therapy. All targets and OARs followed RTOG 0617 contouring guidelines and were directly used in the model validation.

Two plans, PC and LS, were created for each patient using an eight-field static IMRT technique. Standard beam angles based on target laterality were determined from the most common angles used to treat each laterality in the training dataset. All plans were created in the Eclipse (v13.6.23) treatment planning system using the photon optimizer (PO) and anisotropic analytical algorithm (AAA) with a 2.5mm dose grid. Each plan used a single round of optimization of up to 1000 iterations with an intermediate dose calculation for both PC and LS objectives defined in Table 1, with no additional manual optimization. For cases in which the prescription coverage of the PTV exceeded 95% in the clinical plan, the PC

plan was normalized (−5.2% to +2.1%, mean: −0.06%) to match, providing a direct comparison for achieved OAR doses. Otherwise the PC plan was normalized to 95% PTV covered by prescription, the stated planning goal for the trial. In all cases the LS plan was normalized (−2.3% to +2.7%, mean: +0.05%) for the prescription to cover 90% of the PTV, representing the lower limit of acceptable variation, to show maximum achieved lung sparing without violating the protocol. Dose to all other OARs was evaluated after normalization to ensure no protocol violations were created. By providing equal or greater coverage, any improvements observed in the model generated plans can be attributed to the model predictions.

Plan quality at all defined RTOG 0617 dosimetric endpoints was compared between the submitted plans, and both the PC and LS generated plans. As the distribution of data was non-normal, a Sign test (p = 0.05) was used to calculate significance between the paired differences with respect to the clinical plan. The accuracy of the FLM predictions was evaluated at each dosimetric endpoint by determining if the final calculated PC plan was within 1 standard deviation of the predicted value, with the predictive value defined at the midline of the RapidPlan prediction estimate for the specific dose value. Feasibility of utilizing the model-based predictions to determine when a clinical trade-off was necessary prior to optimization was investigated. The predictive threshold was determined using the single institution validation data by setting the achieved Lungs–CTV V_{20Gy} equal to 37% and calculating the associated predictive value at the lower bound of the 95% confidence interval (Figure 2). The accuracy of the predictive threshold was evaluated using the multi-institutional PC generated data, with LS objectives applied when the threshold was exceeded. The "Mixed" dataset in Figure 3 represents combined dosimetric data produced from both the PC and LS optimization objectives, with LS generated plans used for patients that had predictive Lungs-CTV V_{20Gy} greater than the predictive threshold.

Results

Single Institution Model Validation

Treatment plans created for all 9 validation cases using PC objectives produced clinically acceptable results that met RTOG 0617 guidelines. The average 95% confidence interval for the difference between the FLM prediction and PC generated plan for lungs – CTV was ±4.92% for V_{20Gy} and is shown as the separation between the solid and dashed lines in Figure 2. The accuracy of the FLM to predict the dosimetric endpoints for all OARs was evaluated by assessing the linear relationship between the achieved and predicted values, with Lungs-CTV V_{20Gy} represented in Figure 2. The FLM produces an almost ideal predictive-to-achieved relationship across the clinically relevant range, indicative of a highly refined KBP model.

Multi-Institutional RTOG 0617 Model Validation

The RTOG 0617 planning objectives were used to compare plan quality between the 25 multi-institutional clinical plans submitted to RTOG 0617 and the KBP model generated plans (PC, LS, and Mixed). The box-and-whisker plots in Figure 3 illustrate the range of differences between the clinical and KBP plans, with negative values representing improved

sparing produced by the KBP model generated plan. The measure of significance, as determined with the sign test, for all planning objectives in Figure 3 is provided. For the heart, lungs – CTV, and esophagus, the KBP model significantly improved all RTOG 0617 specified dosimetric end-points, with the exception of lungs–CTV V_{20Gy} for the PC and mixed plans, which was not significantly different. There was no significant difference between the clinical and KBP produced plans for maximum dose (volume = 0.03cc) to the spinal cord or brachial plexus. The RTOG 0617 spinal cord planning objective of maximum dose less than 50.5 Gy was achieved for all clinical and KBP produced treatment plans.

The prioritization of achieving Lungs – CTV V_{20Gy} for the RTOG 0617 clinical plans often resulted in unacceptable variations in the PTV coverage goals. Figure 4 compares achieved Lungs – CTV D_{mean} and V_{20Gy} vs. prescription coverage of the PTV for clinical, PC, and LS plans. Both the LS and RTOG 0617 populations had 2 patients with V_{20Gy} exceeding 37%, with 8 RTOG 0617 patients exhibited unacceptable variations in the PTV coverage.

Applying the single institution calculated 95% confidence interval to the linear regression of the multi-institutional predicted versus achieved V_{20Gy} generated the dotted lines in Figure 2. The predictive threshold of V_{20Gy} =31.7% was determined by the intersection of the lower 95% confidence interval at an achieved V_{20Gy} of 37% (Figure 2). Implementing this threshold correctly identifies all 5 cases that exceed $V_{20Gy} > 37\%$ for the PC generated plans and require the LS objectives to reduce the lungs–CTV V_{20Gy} dose. The 95% confidence threshold incorrectly identifies 6 patients that did not exceed $V_{20Gy} > 37\%$ in the PC optimized plans. Utilizing a smaller confidence internal, such as 68%, would result in a predictive threshold of 33.6Gy and correctly identify all 5 cases requiring LS, with only 2 false positives.

Discussion

The variability of plan quality in multi-institutional clinical trials has been investigated in recent years and may impact the conclusions drawn from the outcomes of large cooperative group studies. Implementation of patient specific plan QC tools, including RapidPlan predictions, to minimize this variability and provide anatomy-based justification for acceptable variations in study planning goals is necessary for minimizing the impact of plan quality variability[12,13,15]. The results of this study show that a refined single-institution KBP model can function as a QC tool in a multi-institution clinical trial, significantly improving overall treatment plan quality for most patients. In patients exhibiting difficult anatomy, the KBP model can predict when clinical trade-offs between target coverage and OAR sparing requires a modified set of optimization objective prioritizations to achieve trial defined planning goals.

To effectively implement a KBP model as a QC tool requires a high degree of accuracy for model produced predictions. The overall model accuracy is critical when utilizing these predictions to determine when alternative optimization objectives should be used due to challenging patient anatomy. Previous work by Delaney et al showed inferior quality plans used to created KBP models produced predictions and KBPs with inferior plan quality[16]. The extensive model refinement conducted for this study was completed as a best practice to

ensure the model-generated predictions maximized OAR sparing. The high degree of accuracy for the refined FLM lungs–CTV evaluation point of $V_{20GY} <37\%$ enables the prediction to determine when it would be necessary to accept reduced target coverage to achieve required lung sparing. In this study, a conservatively low prediction threshold of 31.3Gy, representing the lower bound of the single institution 95% confidence interval, was selected to ensure all patients that might exceed $V_{20GY} <37\%$ use LS optimization objectives. Increasing the threshold would reduce the number of cases incorrectly identified as needing LS objectives, and may be achieved by utilizing a larger validation dataset to shrink the 95% confidence interval. The appropriate selection of a refined threshold merits further investigation with a larger patient dataset.

The feasibility of using a refined KBP model developed from a single institution's patient population to improve clinical trial plan quality has been shown in this study. A recent publication by Younge et al provided a similar analysis for a subset of spinal cord SBRT plans submitted to RTOG 0631, showing a single institution KBP model can reduce plan quality variability at study-defined dosimetric endpoints compared to clinical trial submitted plans[17]. This variability of plan quality in multi-institutional studies may limit the efficacy of KBP models generated from patient data created in previous clinical trials. Additionally, a single-institutional model allows for principle investigators to control the quality and consistency of plans and contours used for model generation, ensuring the predictions reflect clinical trade-offs as desired by the study. Finally, a comprehensive, refined single-institution KBP model may require significantly less coordination and effort to construct and validate compared to multi-institutional KBP models.

The FLM was an effective QC tool to improve plan quality for most patients in the multi-institution validation dataset for RTOG 0617 submitted plans. Applying PC optimization objectives significantly improved dosimetric plan quality for most study end points, with the exception of spinal cord D_{max} and lungs–CTV V_{20Gy}. (Figure 3). The PC objectives' inability to improve the lungs–CTV V_{20Gy} was likely due to its high prioritization in the clinical plans, often at the expense of target coverage and sparing of other OARs. This may have impacted the outcomes of RTOG 0617, as Bradley et al hypothesized heart dose may be a primary factor in the decreased survival of the high dose arm of the study[14]. Comparing achieved target coverage to lungs–CTV (Figure 4), it becomes evident that the clinical plans often sacrificed target coverage and increased dose to the heart to achieve V_{20Gy} objectives. When utilizing the prediction threshold to identify whether to use PC or LS objectives, it was possible to achieve target goals while improving lung and heart dose.

The primary limitation of this study is the small patient population used to analyze the impact of KBP on a multi-institutional study. However, we believe observed dosimetric improvements produced from the FLM model may under-represent the dosimetric improvements of the general RTOG 0617 multi-institutional patient population. Eaton et al's secondary analysis of RTOG 0617 documented a significant survival improvement for institutions enrolling a high volume (>4) of patients, possibly due to the high level of experience enabling these institutions to maximize plan quality[13]. All three contributing institutions for this study had high enrollment and produced clinical plans with minimal-to-moderate overall dosimetric improvements. We hypothesis this was especially true for the

heart, which saw modest dosimetric improvements. Results provided in this study support a full secondary review of plan quality improvements for RTOG 0617 to assess the impact of dosimetric plan quality to excessive OAR risk. A second limitation exists in comparing dosimetric improvements produced from a single-institution KBP model to those produced from a multi-institution KBP model. While the presented evidence shows a single-institution KBP model can provide a QC for treatment plan quality in clinical trials, it is beyond the scope of the study to determine the optimal method for model generation. A comparison of plan quality between single- and multi-institutional models would provide guidance regarding which methodology would be most useful for implementation of KBP as a QC tool in future multi-institutional studies.

Conclusion

A single institution KBP model can be applied as a QC tool for multi-institutional clinical trials to improve overall plan quality and provide decision-support to determine the need for anatomy-based dosimetric trade-offs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Moore KL, Brame RS, Low DA & Mutic S Experience-based quality control of clinical intensity-modulated radiotherapy planning. Int. J. Radiat. Oncol. Biol. Phys. 81, 545–551 (2011). [PubMed: 21277097]

2. Nelms BE et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. Pract. Radiat. Oncol. 2, 296–305 (2012). [PubMed: 24674168]

3. Appenzoller LM, Michalski JM, Thorstad WL, Mutic S & Moore KL Predicting dose-volume histograms for organs-at-risk in IMRT planning. Med. Phys. 39, 7446–61 (2012). [PubMed: 23231294]

4. Chanyavanich V, Das SK, Lee WR & Lo JY Knowledge-based IMRT treatment planning for prostate cancer. Med. Phys. 38, 2515–2522 (2011). [PubMed: 21776786]

5. Wu B et al. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. Int. J. Radiat. Oncol. Biol. Phys. 79, 1241–1247 (2011). [PubMed: 20800382]

6. Zhu X et al. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. Med. Phys. 38, 719–726 (2011). [PubMed: 21452709]

7. Li N 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. Biol. Phys. 97, 164–172 (2017). [PubMed: 27979445]

8. Fogliata A et al. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer. Radiat. Oncol. 9, 236 (2014). [PubMed: 25348465]

9. Fogliata A et al. On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer. Radiother. Oncol. 113, 385–391 (2014). [PubMed: 25465726]

10. Tol JP, Delaney AR, Dahele M, Slotman BJ & Verbakel WFAR Evaluation of a knowledge-based planning solution for head and neck cancer. Int. J. Radiat. Oncol. Biol. Phys. 91, 612–620 (2015). [PubMed: 25680603]

11. Fogliata A et al. A broad scope knowledge based model for optimization of VMAT in esophageal cancer: validation and assessment of plan quality among different treatment centers. Radiat. Oncol. 10, 220 (2015). [PubMed: 26521015]

12. Moore KL et al. Quantiying unnecessary normal tissue complication risks due to suboptimal planning : a secondary study on RTOG0126. 92, 228–235 (2016).

13. Eaton BR et al. Institutional Enrollment and Survival among NSCLC Patients Receiving Chemoradiation: NRG Oncology Radiation Therapy Oncology Group (RTOG) 0617. J. Natl. Cancer Inst. 108, 1–8 (2016).

14. Bradley JD et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): A randomised, two-by-two factorial p. Lancet Oncol. 16, 187–199 (2015). [PubMed: 25601342]

15. Tol JP, Dahele M, Delaney AR, Slotman BJ & Verbakel WFAR Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans ? Radiat. Oncol. 1–14 (2015). doi:10.1186/s13014-015-0542-1 [PubMed: 25567003]

16. Delaney AR et al. Effect of Dosimetric Outliers on the Performance of a Commercial Knowledge-Based Planning Solution. 94, (2016).

17. Younge KC et al. Improving Quality and Consistency in NRG Oncology Radiation Therapy Oncology Group 0631 for Spine Radiosurgery via Knowledge-Based Planning. Radiat. Oncol. Biol. 100, 1067–1074 (2018).
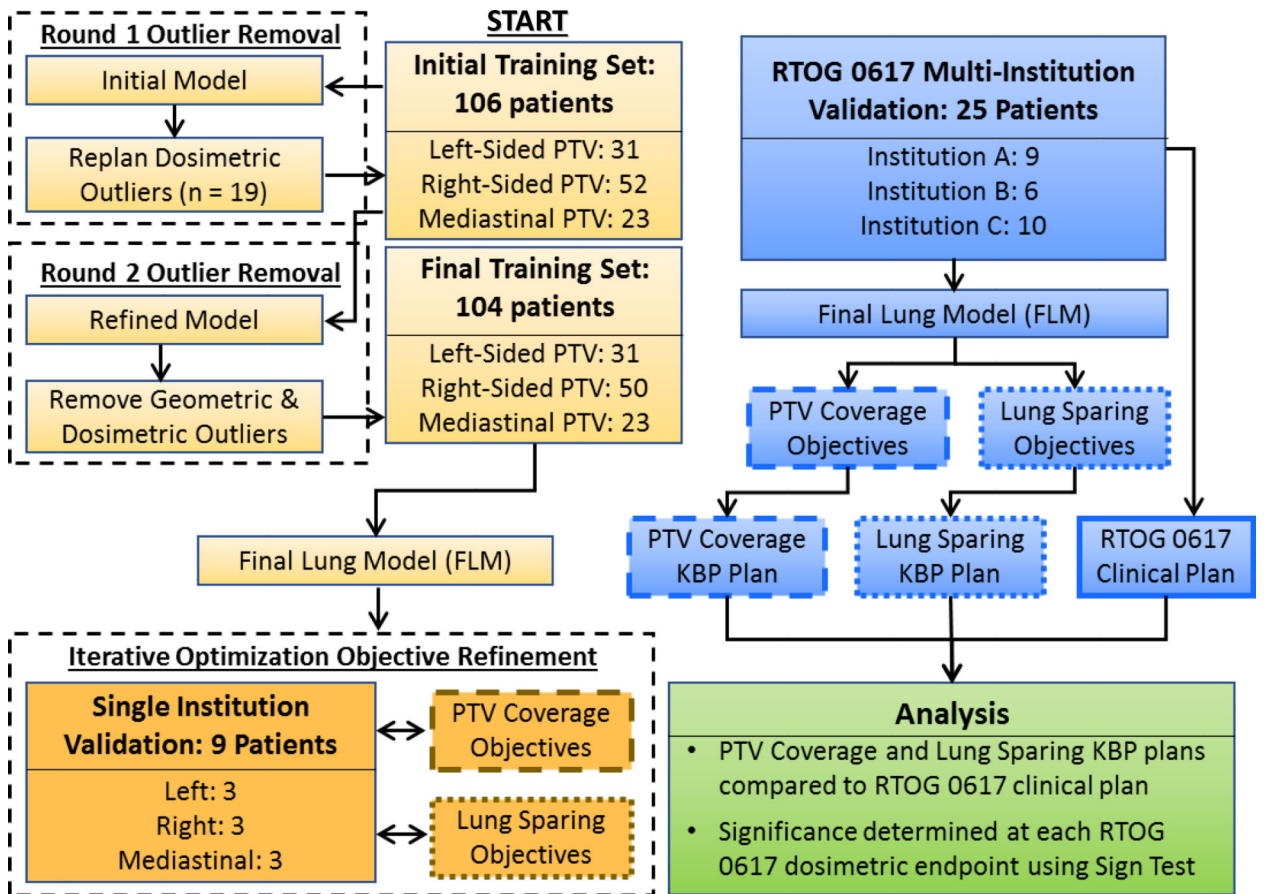
Figure 1:
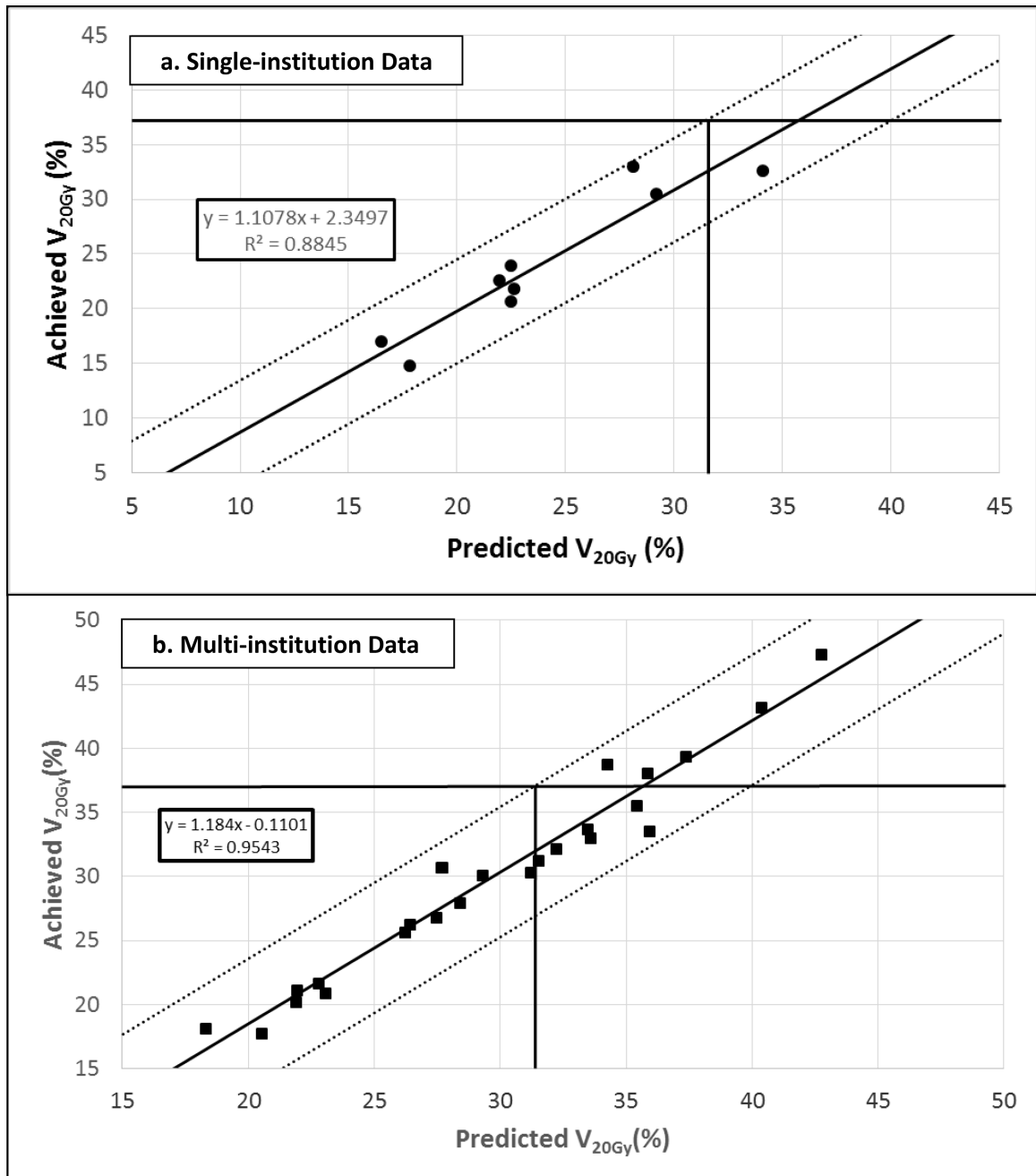Graphical representation of study methodology

Figure 2

Figure 2:

Predictive vs. achieved for the lungs–CTV V_{20Gy} for single- (2a) and multi-institutional (2b) plans using the PC optimization objectives. Dotted lines reflect the 95% confidence intervals determined from the single institution validation. The intersection of the 95% confidence interval at an achieved V_{20Gy} of 37% was used to determine a threshold predictive value, set at 31.3%. Any predictions higher than 31.3% automatically trigger the LS optimization objectives.
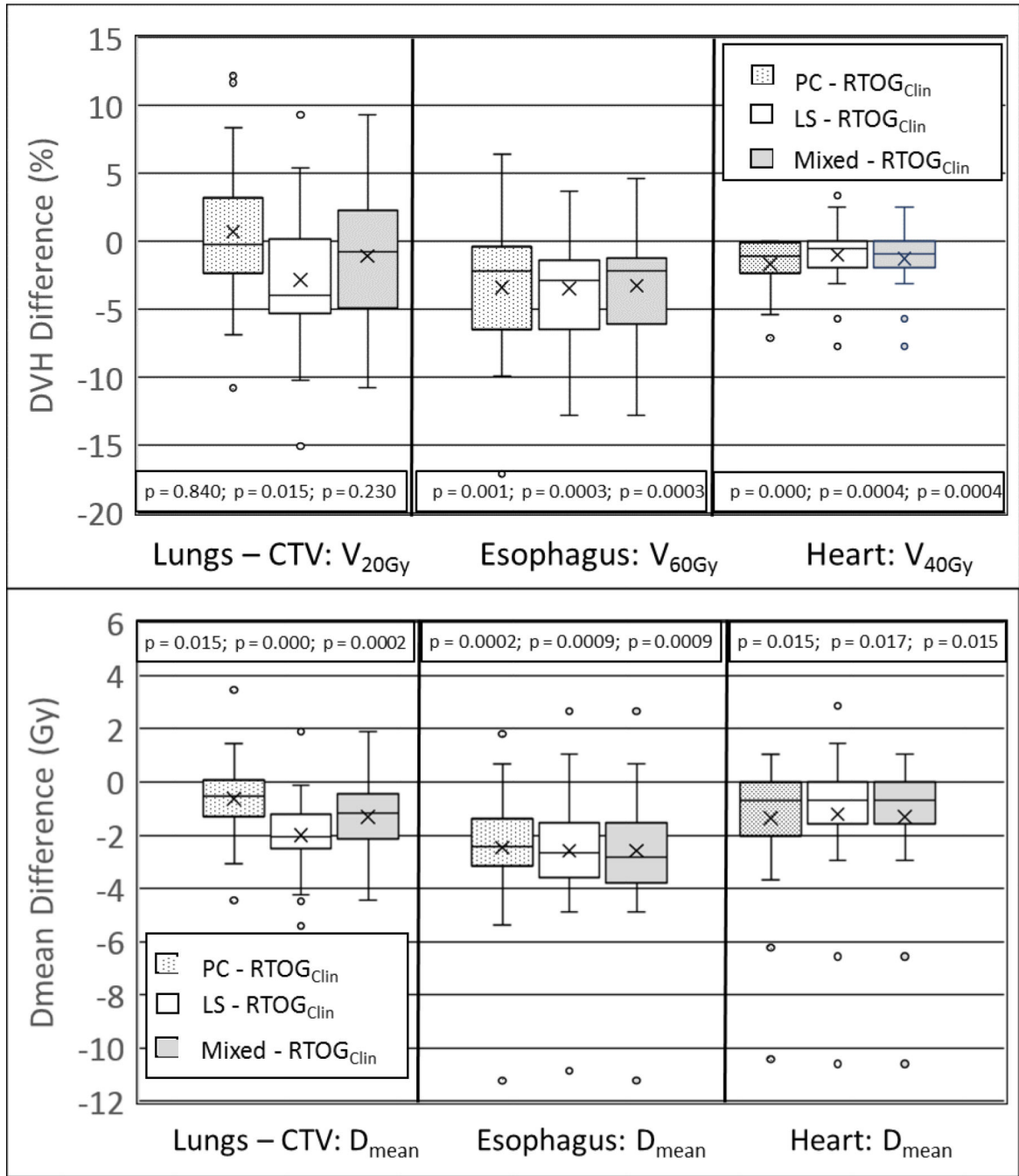
Figure 3:
Dosimetric differences between the PC, LS, and mixed KBP plans and the clinical RTOG 0617 plans at study defined dosimetric endpoints. Negative values indicate lower DVH values in the KBP plans. The outer bounds of the box represent the first and third quartiles, the median is represented by the middle line and the mean value illustrated by the "x". Maximum/minimum values within 1.5 times the inter-quartile range (Q3 – Q1) are represented by the upper/lower error bars, with outliers given by circles.
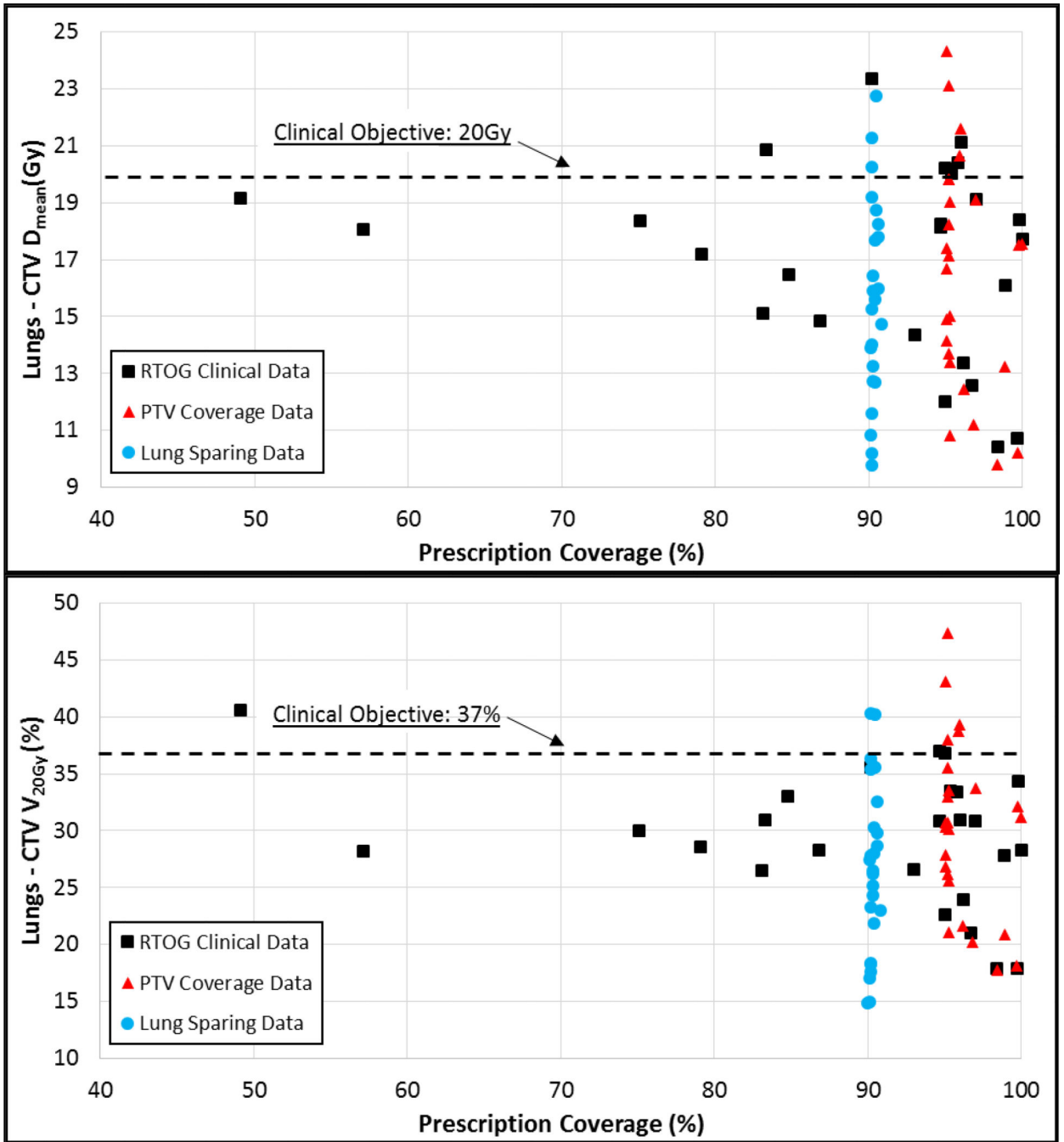
Figure 4:
Lung-CTV sparing (D_{mean} and V_{20Gy}) vs. PTV coverage for the clinical (squares), PTV coverage (triangles), and Lung Sparing (circles) multi-institutional plans.

Table 1:

RTOG 0617/clinical planning goals and optimization objectives for PC and LS plans. The *indicates acceptable variation for clinical PTV coverage of 95% Rx covering 95% of the PTV. Order of prioritization of clinical planning goals was spinal cord, brachial plexus, lungs, heart, PTV, and esophagus. Lower/upper objectives reduce/increase dose to the specified level.

Targets	RTOG/Clinical Planning Objectives			Knowledge-Based Optimization Objectives				
	Objective	Specified Goal	Variation	Type	Volume	Dose	PC Weight	LS Weight
CTV	N/A			Lower	100%	102%/101%	110	110
PTV	100% Rx	95% / 95%	90% / *	Upper	0%	107%	110	100
	Dmax	120% / 120%	125%	Lower	100%	101%/100%	110	110
Inner Ring	N/A			Lower	100%	102%/100%	110	110
Organs-at-Risk	RTOG/Clinical Planning Objectives			Knowledge-Based Optimization Objectives				
	Objective	Specified Goal	Variation	Type	Volume	Dose	PC Weight	LS Weight
Lungs - CTV	Dmean	20 Gy / 20 Gy	N/A	Upper	Gen.	3000 cGy	110	110
	20 Gy	37% / 37%	N/A			2000 cGy	110	130
	5 Gy	N/A / 60%	N/A			500 cGy	110	110
	N/A			Line	Gen.	Gen.	70	90
Lung_Lt	N/A			Line	Gen.	Gen.	50	75
Lung_Rt	N/A			Line	Gen.	Gen.	50	75
Esophagus	Dmean	34 Gy / 34 Gy	N/A	Upper	0%	103%	75	60
	60 Gy	N/A / 17%	N/A		Gen.	6000 cGy	75	60
	N/A			Line	Gen.	Gen.	50	50
Heart	60 Gy	33% / 25%	N/A	Upper	0%	103%	90	75
	45 Gy	66%/35%	N/A		Gen.	5000 cGy	90	75
	40 Gy	100% / N/A	N/A		Gen.	3000 cGy	90	75
	30 Gy	N/A / 50%	N/A	Line	Gen.	Gen.	50	50
Liver	N/A			Line	Gen.	Gen.	40	40
Spinal Cord	Dmax	50.5 Gy / 45 Gy	N/A	Upper	0%	4300 cGy	130	130
				Line	Gen.	Gen.	50	50
Spinal Cord+5mm	Dmax	N/A / 50 Gy	N/A	Upper	0%	4700 cGy	130	130
				Line	Gen.	Gen.	50	50
Brachial Plexus	Dmax	N/A / 66 Gy	N/A	Upper	0%	6400 cGy	120	100