

# Using Novel Statistical Techniques to Accurately Determine the Predictive Dose Range in a Study of Overall Survival after Definitive Radiotherapy for Stage III Non-Small Cell Lung Cancer in Association with Heart Dose

Joshua R. Niska<sup>1</sup>, Jiuyun Hu<sup>2</sup>, Jing Li<sup>3</sup>, Michael G. Herman<sup>4</sup>, Cameron S. Thorpe<sup>5</sup>, Steven E. Schild<sup>5</sup>, Mirek Fatyga<sup>5\*</sup>

<sup>1</sup>Department of Radiation Oncology, Banner MD Anderson Cancer Center, Phoenix, Arizona, USA

<sup>2</sup>School of Computing, Informatics, and Decision Systems Engineering, Tempe, Arizona, USA

<sup>3</sup>School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

<sup>4</sup>Department of Radiation Oncology, Mayo Clinic, Rochester, Minnesota, USA

<sup>5</sup>Department of Radiation Oncology, Mayo Clinic, Phoenix, Arizona, USA

Email: Joshua.Niska@bannerhealth.com, jiuyunhu@asu.edu, jli3175@gatech.edu, Herman.Michael@mayo.edu, Thorpe.Cameron@mayo.edu, sschild@mayo.edu, \*fatyga.mirek@mayo.edu

**How to cite this paper:** Niska, J.R., Hu, J.Y., Li, J., Herman, M.G., Thorpe, C.S., Schild, S.E. and Fatyga, M. (2021) Accurately Determine the Predictive Dose Range in a Study of Overall Survival after Definitive Radiotherapy for Stage III Non-Small Cell Lung Cancer in Association with Heart Dose *Journal of Cancer Therapy*, 12, 505-529. <https://doi.org/10.4236/jct.2021.129044>

**Received:** August 14, 2021

**Accepted:** September 26, 2021

**Published:** September 29, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

**Purpose:** Recent studies of radiotherapy (RT) for stage III non-small-cell lung cancer (NSCLC) have associated high dose to the heart with cardiac toxicity and decreased overall survival (OS). We used advanced statistical techniques to account for correlations between dosimetric variables and more accurately determine the range of heart doses which are associated with reduced OS in patients receiving RT for stage III NSCLC. **Methods:** From 2006 to 2013, 119 patients with stage III NSCLC received definitive RT at our institution. OS data was obtained from institutional tumor registry. We used multivariate Cox model to determine patient specific covariates predictive for reduced overall survival. We examined age, prescription dose, mean lung dose, lung V20, RT technique, stage, chemotherapy, tumor laterality, tumor volume, and tumor site as candidate covariates. We subsequently used novel statistical techniques within multivariate Cox model to systematically search the whole heart dose-volume histogram (DVH) for dose parameters associated with OS. **Results:** Patients were followed until death or 2.5 to 81.2 months (median 30.4 months) in those alive at last follow up. On multivariate analysis of whole heart DVH, the dose of 51 Gy was identified as a threshold dose above which the dose volume relationship becomes predictive for OS.

We identified V55Gy (percentage of the whole heart volume receiving at least 55 Gy) as the best single DVH index which can be used to set treatment optimization constraints (Hazard Ratio = 1.044 per 1% increase in heart volume exposed to at least 55 Gy,  $P = 0.03$ ). Additional characteristics correlated with OS on multivariate analysis were age, stage (IIIA/IIIB), and administration of chemotherapy. **Conclusion:** Doses above 51 Gy, applied to small volumes of the heart, are associated with worse OS in stage III NSCLC patients treated with definitive RT. Higher stage, older age and lack of chemotherapy were also associated with reduced OS.

## Keywords

Lung Cancer, Cardiac Toxicity, Lung Radiation Therapy, Non-Small Cell Lung Cancer, Radiation Toxicity

---

## 1. Introduction

Late cardiac effects years to decades after thoracic radiotherapy (RT) have been well-described [1]. However, recent prospective studies in lung cancer have identified radiation-related cardiac events occurring on an earlier timeframe of months to years. Heart dose volume histogram (DVH) variables associated with cardiac toxicity in these prospective studies include: mean heart dose (MHD) [2] [3] [4], heart V5 (volume of the heart receiving at least 5 Gy) [2] [3] [4] [5], heart V30 [2] [3], heart V35 [6], heart V55 [4], heart V60 [5], mean left ventricle dose [3] [5], left ventricle V5 and V30 [3] [5], mean left atrium dose [5], left atrium V30 [5], and right atrium V60 [5].

Heart dose has not only been associated with cardiac events [2] [3] [4] [5] [6] in prospectively evaluated patients but also overall survival (OS) [7] [8] [9]. In contrast to expectations, the Radiation Therapy Oncology Group (RTOG) 0617 trial demonstrated decreased OS in patients randomized to higher dose radiotherapy for locally advanced non-small-cell lung cancer (NSCLC) [7]. Subsequent analysis of RTOG 0617 associated increased heart V40 [8] and heart V50 [10] with decreased OS.

Prior published studies determined a wide range of associations between dosimetric variables and OS which made their consistent clinical application difficult and even raised doubts as to the veracity of the findings [11]. The inconsistencies among studies are most likely attributable to correlations among Dose Volume Histogram (DVH) variables which were not accounted for in a conventional statistical analysis. To address these shortcomings, we used advanced statistical techniques to systematically search for heart DVH variables which were most predictive for the OS in a cohort of 119 patients, treated for stage III NSCLC with definitive RT at Mayo Clinic Arizona (MCA). We selected well established statistical techniques which are applicable to highly correlated covariates but also enhanced them with novel constraints which reflect radiobiologi-

cal knowledge specific to RT. These additional constraints improve generalizability of the model and make the results easier to interpret.

The increasing use of intensity modulated RT (IMRT) and proton beam therapy (PBT) may allow for more precise cardiac sparing radiation plans, provided that evidence based treatment planning constraints on heart dosimetry can be reliably established.

## 2. Methods and Materials

### 2.1. Patient Characteristics

From 2006 to 2013 at Mayo Clinic Arizona, 119 stage III NSCLC patients were treated with definitive RT. RT was delivered using involved-field technique and either 3-dimensional conformal RT (3D-CRT) or IMRT. Patient characteristics are shown in **Table 1**.

**Table 1.** Baseline patient and treatment characteristics. Treatments were conventionally fractionated at 1.8 Gy - 2.0 Gy per fraction.

	N (%)	Median (Range)
Follow-Up of the Surviving Patients (months)		30.4 (2.5 - 81.2)
Follow-Up for All Patients (months)		18 (1.2 - 81.2)
Age (years)		70.5 (41.7 - 91.1)
Prescription Dose (Gy)		62 (43.1 - 74.0) 1.8
Mean Lung Dose (Gy)		13.6 (4.4 - 22.4)
Lung V20 (%)		23.9 (5.7 - 41.5)
Tumor Volume [cc] (CTV)		118.5 (1.1 - 706)
Technique 3D-CRT	49 (41.2%)	
IMRT	70 (58.8%)	
Stage IIIA	72 (60.5%)	
IIIB	47 (39.5%)	
Chemo Yes	106 (89.1%)	
No	13 (10.9%)	
Laterality Left	44 (37%)	
Right	74 (62.2%)	
Undefined	1 (0.8%)	
Site Lower Lobe	26 (21.8%)	
Middle Lobe	5 (4.2%)	
Upper Lobe	80 (67.2%)	
Bronchus	5 (4.2%)	
Undefined	3 (2.5%)	
Alive at Last Follow-up	47 (39.5%)	
Deceased at Last Follow-Up	72 (60.5%)	

## 2.2. Treatment Planning Heart Constraints

Typical dose constraints on the whole heart structure during treatment planning were: Maximum dose < 62 Gy, Mean dose < 26 Gy,  $V_{30\text{Gy}} < 46\%$  ;  $V_{40\text{Gy}} < 33\%$ .

## 2.3. Heart Dose Extraction

The whole heart was contoured for each patient on the radiation planning computed tomography (CT) image using Eclipse treatment planning system (Varian, Inc). Dosimetric information was extracted from Eclipse using the Eclipse Application Programmer Interface (ESAPI) software and reprocessed for statistical analysis using proprietary institutional software. The typical planning image was acquired with axial spacing of 2 mm and 1 - 2 mm voxels in the anterior-posterior and medial-lateral dimensions. Dose extraction was performed using a rectangular grid with the dimensions of the planning image.

## 2.4. Dosimetric Analysis

The dose to the heart was quantified using dosimetric index  $V_D$  which was defined as the percentage of the volume of the heart (or heart segment) receiving dose  $\geq D$  in Gy. For each dose-volume histogram (DVH), a range of  $V_D$  indices was extracted, with dose  $D$  varying in 1 Gy steps between 5 Gy and 60 Gy. The dose was not converted to biologically equivalent dose as all treatments were conventionally fractionated photons at 1.8 - 2.0 Gy per fraction.

## 2.5. Statistical Analysis

Possible association between heart dosimetry and OS is most commonly investigated using the multivariate Cox model with heart dosimetry represented by preselected heart DVH variables (e.g., a  $V_D$  which is the percentage of heart volume receiving dose  $D$ , or greater). The drawback of this approach is that p-values associated with DVH variables need to be scaled by False Discovery Rate (FDR) correction. Since the dosimetric variables are highly correlated, the FDR correction can be overly strict and therefore may find none of the dosimetric variables being significant, particularly in studies with limited patient numbers. Also, this approach cannot assess the joint effect of dosimetric variables on OS. To address these limitations a multivariate Cox model can be adopted in which heart dosimetry is represented by the linear combination of DVH features, *i.e.*,

$$h(t) = h_0(t) \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p),$$

which links the hazard function  $h(t)$  with dosimetric variables  $X_1, \dots, X_p$ . For example,  $X_1, \dots, X_p$  can be  $V_{5\text{Gy}}, V_{6\text{Gy}}, \dots, V_{59\text{Gy}}, V_{60\text{Gy}}$ . The conventional multivariate Cox model may suffer from multicollinearity due to the high correlation between the dosimetric variables. Another challenge is the curse of dimensionality as the number of dosimetric variables included in the multivariate model is typically quite large compared with the limited sample size. To address these is-

sues, we took advantage of modern developments in statistical analysis by adding constraints on the coefficient estimates, which are known as the “variable selection techniques” [12]. The well-known Lasso model [13] adds an L1 penalty on the coefficient estimates, *i.e.*,  $\sum_{i=1}^p |\beta_i| < s$ , which has the effect of suppressing the small-effect coefficients to be zero and thus selecting the subset of important dosimetric variables simultaneously. To further account for the high correlation among the dosimetric variables, the fused Lasso model [14] includes an additional L1 penalty on dosimetric variables with adjacent dose levels *i.e.*,  $\sum_{i=1}^p |\beta_i - \beta_{i-1}| < \gamma$ . The upper bounds in these constraints,  $s$  and  $\gamma$ , are selected using a grid search to optimize a commonly used model selection criterion. Based on the work of Dai and Breheny [15] we use leaveoneout cross validation of linear predictors during the grid search to find parameters  $s$  and  $\gamma$  associated with the lowest cross validation error (**Supplement S2.2**).

In this paper, we adopted the fused Lasso as our base formulation but added two additional constraints inspired by biomedical domain knowledge. The resulting model is called knowledge-constrained Lasso (KC-Lasso). The two constraints are non-negativity and monotonicity of coefficients for dosimetric variables. The non-negativity means that we require  $\beta_i \geq 0$ ,  $i = 1, \dots, p$ . This constraint reflects a biologically motivated hypothesis that increasing  $V_D$  poses either a higher hazard risk or no significant risk but cannot lower the risk. The monotonicity constraint specifies that  $\beta_1 \leq \beta_2 \leq \dots \leq \beta_p$  where  $\beta_1$  to  $\beta_p$  are coefficients corresponding to  $V_D$ 's with  $D$  from the lowest to the highest dose levels and is motivated by radiobiology. If the same volume is irradiated to a higher dose, the hazard ratio associated with the irradiation cannot decrease since higher doses are always associated with lower cell survival fractions. Lower cell survival fraction may keep clinical toxicity the same or make it worse, but it cannot make it better (**Supplement S2.1**). Integrating the non-negativity and monotonicity constraints into the fused Lasso formulation, the resulting KC-Lasso model can be estimated by the “penalized” function in the R package.

The analysis was performed in two distinct steps: In the first step we selected patient-specific covariates which were predictive for the OS, using Cox model with Akaike Information Criterion (AIC). Following covariates were considered: stage, chemotherapy, age, prescription dose, mean lung dose, lung V20, tumor site, and laterality. Tumor volume was not included as a candidate covariate because of strong correlation between volume and stage (**Supplement S2.5**). In the second step we retained only predictive patient specific covariates and additionally included dosimetric covariates, without penalizing patient specific covariates.

We applied the KC-Lasso model with two different resolutions, one with 1 Gy spacing between indices, *i.e.*, 5 Gy, 6 Gy, ..., 59 Gy, 60 Gy included in the model, which we called the “dense fit”; and the other with 5 Gy spacing, *i.e.*, 5 Gy, 10 Gy, ..., 55 Gy, 60 Gy, which we called the “sparse fit”. The selected patient-specific covariates were included in each model. The specific purpose of

varying the dose step was to test the self-consistency of the approach as the step size selection is arbitrary and the final evaluation of the Hazard Ratios should not depend strongly on the step size selection. Greater resolution provides a better estimate of the effective dose threshold and the final Hazard Ratio, at a cost of greater complexity of the model in its final applications.

The KC-Lasso model can identify the dose threshold (or thresholds) beyond which the dose volume variables,  $V_{D \geq D^*}$  become predictive for OS. However, the model itself cannot be used directly in commercially available optimization packages which set constraints on individual indices only. For this reason we also fit a family of multivariate Cox models, each with only one  $V_D$  index representing heart dosimetry, to obtain a (slightly less precise) model which is directly usable as a dose constraint in commercially available treatment planning systems. Since KC-Lasso identifies the range of doses which are predictive for OS, we did not scale p-values in the simplified approach by the FDR correction if they fell within the range indicated by KC-Lasso. In clinical applications the single index constraint can be used to set the optimization constraint based on limiting the Hazard Ratio, while the full formulation of KC-Lasso can be used for final evaluation of the Hazard Ratio in the plan.

KC-Lasso is not the only statistical methodology that could be applied to DVH analysis. The alternatives could include Lasso and Fused Lasso methods without knowledge constraints or entirely different statistical approaches which account for correlations among indices. To explore potential alternatives, we applied Lasso, Fused Lasso, and Elastic Net [16] models to our data set and compared the results to KC-Lasso.

### 3. Results

Median follow-up for all patients was 18 months (range 1.1 to 81.2 months). At last follow-up, 47 patients (39.5%) were alive (**Table 1**). Median follow-up of the patients alive at last follow up evaluation was 30.4 months (range 2.5 to 81.2 months). Three patient-specific variables were associated with OS in all models: age before RT, disease stage, and receipt of chemotherapy. Prescription dose, mean lung dose, lung V20, radiation technique (3D-CRT vs IMRT), timing of chemotherapy (concurrent vs sequential), tumor laterality, tumor site and volume were not significant.

#### 3.1. Patient Specific Covariates

Overall Survival was worse for older patients (HR = 1.04 per year, P = 0.01), worse for stage IIIB vs IIIA (HR = 1.78, P = 0.02), and better for patients who received chemotherapy (HR = 0.46, P = 0.04).

#### 3.2. Whole Heart Dosimetry Using KC-Lasso Multivariate Analysis

KC-lasso identified a dose threshold,  $D^*$ , above which the dose volume variables,  $V_{D \geq D^*}$ , become predictive for OS. In the dense fit,  $D^* = 51$  Gy, *i.e.*, all the coeffi-

coefficients corresponding to dosimetric variables with  $D < 51$  Gy are zero while the first non-zero coefficient occurs for  $V_{51}$  and all the coefficients with  $D \geq 51$  Gy are non-zero. In the sparse fit,  $D = 55$  Gy. Both results are summarized in **Table 2**.

### 3.3. Whole Heart Dosimetry Using Single $V_D$ Index

**Table 3** shows p-values associated with  $V_D$  indices obtained by fitting a family of Cox models, each using a single  $V_D$  index to represent heart dosimetry, spaced in 5 Gy increments from V5 to V60. Heart V55 predicted OS in a statistically significant manner, while V50 and V60 were nearly statistically significant. The hazard ratios (HRs) for OS are worse for increasing heart V55 (HR 1.044 per 1% increase in heart volume exposed to at least 55 Gy,  $P = 0.03$ ) (**Table 4**).

### 3.4. KC-Lasso Consistency Check

The two variants of KC-Lasso (**Table 2**) and a conventional model (**Table 4**) provide a consistency check on the estimates of Hazard Ratio (HR) with both approaches. The ratio between coefficients in “dense” and “sparse” fit of KC-Lasso (**Table 2**) is approximately 1:5, which is the same as the step size ratio, hence both models will evaluate to a similar HR value. Similarly, the value of the coefficient in a conventional model is 0.043 (**Table 4** and **Table 1S**), which is

**Table 2.** Summary of coefficients for  $V_D$  features which are predictive for OS in the KC-Lasso model. Results are shown for two models, each using an array of  $V_D$  indices as input. In the “dense” model indices are spaced by 1 Gy, whereas in the “sparse” model indices are spaced by 5 Gy. For both models, coefficients are zero in the V1 - V50 range, and increase with dose thereafter. The mathematical formula needed to calculate the hazard ratio is shown at the bottom of the table. The p-value associated with dosimetric variables is shown in the last column. Note that weights in the “dense” model are approximately 5 times lower than the weights in the “sparse” model, which means that weights scale in proportion to the dose step. Both models will evaluate to a similar Hazard Ratio, showing the consistency between the two models. Summary of KC-Lasso DVH feature.

DVH index	V1 - V50	V51	V52	V53	V54	V55	V56	V57	V58	V59	V60	p-value
$\beta_{dense}$	0.0	0.003	0.0031	0.0033	0.0035	0.0038	0.004	0.0046	0.005	0.0055	0.006	0.019
$\beta_{sparse}$	0.0	--	--	--	--	0.021	--	--	--	--	0.033	0.02

KC-Lasso model hazard ratio =  $e^{\sum \beta_i (V_{-D})^i}$

**Table 3.** A summary of P-values associated with DVH indices for a family of Cox models that represent heart dosimetry as a single, whole heart DVH. Index  $V_D$  indicates percentage of heart volume receiving a dose greater or equal to  $D$ [Gy]. P-values are lowest in the same range of  $V_D$  as non-zero indices of the KC-Lasso model. Since the lowest p-value is associated with V55, which is also located in the middle of the KC-Lasso range, we choose V55 as the best approximation which can be used in treatment planning. The Hazard Ratio associated with V55 is HR = 1.044 for each 1% of the heart volume exposed to at least 55 Gy.

<b>DVH index</b>	<b>V5</b>	<b>V10</b>	<b>V15</b>	<b>V20</b>	<b>V25</b>	<b>V30</b>
P-value	0.55	0.53	0.46	0.37	0.31	0.42
<b>DVH index</b>	<b>V35</b>	<b>V40</b>	<b>V45</b>	<b>V50</b>	<b>V55</b>	<b>V60</b>
P-value	1.00	0.65	0.33	0.09	<b>0.03</b>	0.08

**Table 4.** Multivariate Cox model for OS using whole heart V55 as a single DVH index representing heart dosimetry. V55 represents percentage of whole heart volume receiving dose 55 Gy, or greater. The Hazard Ratio for cardiac toxicity can be calculated as  $HR_{cardiac} = e^{0.0439V55} \cong (1.044)^{V55}$ . Equations and examples needed to calculate an individualized HR for OS using a specific patient's variables are provided in the Supplement.

Model	Heart V55	
	HR	P-value
$V_D$ (per 1%)	1.044	0.03
Age (per year)	1.04	0.01
Stage IIIB	1.78	0.02
Chemotherapy (Concurrent or Sequential)	0.46	0.04

comparable to coefficients in the “sparse” KC-lasso fit (**Table 2**), and will thus yield a similar estimate of HR. All models are approximations and one does not expect an exact agreement among them, but one does expect a reasonable consistency of HR estimates.

### 3.5. Alternative Statistical Approaches

The three alternative statistical approaches (Lasso and Fused Lasso without knowledge constraints and Elastic Net without knowledge constraints) generated fits to data which were of comparable statistical significance to KC-Lasso but each of the three approaches had features which were difficult to interpret, like negative correlation coefficients or isolated correlation coefficients at a single dose. Hence knowledge based constraints, similar to constraints in KC-Lasso, are likely needed to create models which are both intuitively understandable and more likely to be generalizable to other data sets. A detailed discussion of the comparisons among competing statistical techniques can be found in the supplemental section (**Supplement S2.4**).

## 4. Discussion

We used advanced statistical techniques to overcome limitations of conventional statistical methods which are often used to search for associations between heart dosimetry and OS in lung cancer patients. Conventional analyses use the Cox model with preselected DVH variables (in a univariate or multivariate setting) and seek to establish statistically significant associations between DVH variables and OS. These approaches raise False Discovery (FDR) concerns and ignore strong correlations between DVH variables, which can lead to variable results when studies are compared (**Table 5**) [11]. Additional discussion of the limitations of the univariate approach is provided in the supplemental section (**Supplement S2.3**).

The model introduced in our work (KC-Lasso) treats the entire DVH as input and finds a contiguous range of DVH variables predictive for OS (the sensitivity

**Table 5.** Studies published since 2008 reporting an association between survival outcomes and heart dose.

Study	N	Follow-up	Type of Cancer	RT Technique (s)	RT Prescription	Major Finding (s)
Odense Univ. <sup>f</sup> Schytte, 2010	250	7.9 yrs <sup>^</sup>	NSCLC	3D-CRT	60 - 80 Gy	LV mean $\geq$ 14.5 Gy $\rightarrow$ $\downarrow$ OS (p = 0.06)
Euro2K <sup>t</sup> Tukenova, 2010	4122	26 yrs <sup>^</sup>	Pediatric	2D-RT	NR	Heart mean 1 Gy $\rightarrow$ $\uparrow$ cardiac death excess RR 60% Heart mean 1-4.9 $\rightarrow$ $\uparrow$ cardiac death RR 2.5 Heart mean 5-14.9 Gy $\rightarrow$ $\uparrow$ cardiac death RR 12.5 Heart mean $\geq$ 15 Gy $\rightarrow$ $\uparrow$ cardiac death RR 25.1
RTOG 0617 <sup>a</sup> Bradley, 2015	544	1.9 yrs <sup>^</sup>	NSCLC	3D-CRT (51.5%) IMRT (48.5%)	60 Gy (57.5%) 74 Gy (42.5%)	Heart V5 $\rightarrow$ $\downarrow$ OS HR 1.007 per 1% Heart V30 $\rightarrow$ $\downarrow$ OS
RTOG 0617 <sup>a</sup> Eaton, 2016	495	NR	NSCLC	3D-CRT (51.7%) IMRT (48.3%)	60 Gy (58.2%) 74 Gy (41.8%)	Heart V50 $\rightarrow$ $\uparrow$ Grade 5 AE
RTOG 0617 <sup>a</sup> Chun, 2017	482	1.8 yrs <sup>^</sup>	NSCLC	3D-CRT (53%) IMRT (47%)	60 Gy (58%) 74 Gy (42%)	Heart V40 $\rightarrow$ $\downarrow$ OS HR 1.01 per 1%
William Beaumont <sup>t</sup> Johnson, 2017	178	1.4 yrs <sup>^</sup>	NSCLC	3D-CRT (38.4%) IMRT (61.6%)	64 Gy <sup>^</sup>	Heart V30 $\rightarrow$ $\downarrow$ OS HR 1.013 per 1%
Univ. of Manchester <sup>t</sup> McWilliam, 2017	1101	3 - 36 mos	NSCLC	3D-CRT IMRT SBRT (7.4%)	55 Gy (non-SBRT) 60 Gy in 5 fx (SBRT)	Heart base mean > 16.3 Gy $\rightarrow$ $\downarrow$ OS HR 1.25 (non-SBRT) Heart base mean > 8.5 Gy vs < 8.5 Gy $\rightarrow$ $\downarrow$ OS (non-SBRT) Heart base mean > 6.3 Gy $\rightarrow$ $\downarrow$ OS HR 2.11 (SBRT)
Mayo Clinic PCI <sup>f</sup> Sio, 2017	76	5.5 yrs <sup>^</sup>	Thoracic -Breast (57%) -NSCLC (17%) -Upper GI (11%)	2D-RT 3D-CRT IMRT (1.3%) SBRT (2.6%)	53.4 Gy <sup>*</sup>	Heart mean $\rightarrow$ $\downarrow$ OS HR 2.01 per 1 Gy Heart mean $\rightarrow$ $\uparrow$ non-cancer death HR 1.49 per 1 Gy Heart max $\rightarrow$ $\downarrow$ OS after PCI HR 1.02 per 1 Gy Heart V40 $\rightarrow$ $\uparrow$ non-cancer death HR 1.32 per 1%
Washington Univ. <sup>f</sup> Speirs, 2017	322	1.2 yrs <sup>^</sup>	NSCLC	3D-CRT (60%) IMRT (40%)	66 Gy <sup>^</sup>	Heart V50 $\rightarrow$ $\downarrow$ OS HR 1.23 per 1% Heart V55 $\rightarrow$ $\downarrow$ OS HR 1.85 per 1% (CRT only)
Multicenter <sup>t</sup> Stam, 2017	803	2.9 yrs <sup>^</sup>	NSCLC	SBRT	54 Gy in 3 fx <sup>^</sup>	LA max $\rightarrow$ $\uparrow$ non-cancer death HR 1.005 per 1 Gy SVC D90 $\rightarrow$ $\uparrow$ non-cancer death HR 1.025 per 1 Gy
Meta-Analysis <sup>a</sup> Taylor, 2017	40,781	10 yrs <sup>^</sup>	Breast	2D-RT	NR	Heart mean $\rightarrow$ $\uparrow$ cardiac death RR 1.04 per 1 Gy
IDEAL-CRT <sup>a</sup> Vivekanandan, 2017	78	35 mos <sup>^</sup>	NSCLC	3D-CRT ("most") IMRT ("some")	67.7 Gy <sup>*</sup>	LA wall V63 > 2.2% $\rightarrow$ $\downarrow$ OS HR 1.52
Washington Univ. <sup>f</sup> Contreras, 2018	400	17 mos <sup>^</sup>	NSCLC	3D-CRT (59%) IMRT (41%)	66 Gy <sup>^</sup>	Heart V50 $\rightarrow$ $\downarrow$ OS HR 1.02 per 1%

**Continued**

Princess Margaret <sup>†</sup> Wong, 2018	189	35.3 mos <sup>^</sup>	NSCLC	SBRT	48 Gy in 4 fx (47.1%) 54 - 60 Gy in 3 fx (24.3%)	Ventricle max → ↑non-cancer death HR 1.02 per 1 Gy
Multicenter <sup>a</sup> Xue, 2019	94	58 mos <sup>^</sup>	NSCLC	3D-CRT	70 Gy <sup>^</sup>	Pericardium V30 > 29% → ↓OS (HR 1.019 per 1%) Pericardium V55 > 21% → ↓OS (HR 1.03 per 1%)
Dana-Farber Atkins 2019	748	20.4 mos <sup>^</sup>	NSCLC	3D-CRT (78.1%) IMRT (21.9%)	66 Gy <sup>^</sup>	Heart mean → ↑ACM HR 1.02 per 1 Gy With pre-existing CHD, heart mean ≥ 10 Gy → ↑ACM HR 1.34 Without pre-existing CHD, no association between heart mean and ACM
<b>Present Study<sup>†</sup></b> Niska 2019	119	1.5 yrs <sup>^</sup>	NSCLC	3D-CRT (41.2%) IMRT (58.8%)	62 Gy <sup>^</sup>	Heart V55 → ↓OS HR 1.044 per 1% Heart V51 - V60 range predictive for OS

Abbreviations: 2D-RT, 2-dimensional RT; 3D-CRT, 3-dimensional conformal RT; ACM, all-cause mortality; AE, adverse event; CHD, coronary heart disease; D(xx), minimum dose to xx% of the volume; fx, fractions; GI, gastrointestinal; Gy, Gray; HR, hazard ratio; IMRT, intensity-modulate RT; LA, left atrium; LV, left ventricle; mos, months; NR, not reported; NSCLC, non-small cell lung cancer; OS, overall survival; PCI, percutaneous coronary intervention; RR, relative risk; RT, radiotherapy; RTOG, Radiation Therapy Oncology Group; SBRT, stereotactic body RT; SVC, superior vena cava; univ., university; V(x), volume receiving at least x Gy; yrs, years. <sup>^</sup>indicates median <sup>\*</sup>indicates mean <sup>\*</sup>indicates prospective <sup>†</sup>indicates retrospective.

range). However, the model itself cannot be used directly in commercially available optimization packages which set thresholds on individual indices only. Hence, we supplemented our analysis with a conventional approach, which used a single DVH variable to represent heart dosimetry in a multivariate Cox model and searched for the model in which this variable had the greatest statistical significance (**Table 3** and **Table 4**). We argue that the variable with greatest statistical significance can be selected without concerns for FDR, as long as it belongs to the “sensitivity range” selected by the KC-Lasso model. Clinically, the conventional approach would be used to establish optimization constraints, by setting a limit on the Hazard Ratio, while the more complete KC-Lasso model could be used to evaluate the Hazard Ratio in the treatment plan. A more detailed discussion of the limitations of the conventional model can be found in the supplemental section (**Supplement S2.3**).

Our findings build upon prior studies that also show an association between high doses to relatively small volumes of the heart and decreased survival in NSCLC [7] [8] [10] [17]-[24]. Consistent with the findings of other investigators, our model predicts worse OS for older age [25], more advanced stage [26], and lack of chemotherapy [27].

Hazard ratios for OS associated with heart irradiation in our study are of comparable magnitude to the HRs for older age. Using the conventional model approximation as an illustration, each additional 1% of heart receiving ≥55 Gy carries similar OS impact to an additional year of older age (**Table 4**).

**Table 5** provides a summary of the existing literature associating heart dose with OS [4] [7] [8] [10] [17]-[24] [28] [29] [30] [31] [32]. While our study found a range of doses for which DVH variables were associated with OS (V51 - V60),

other studies have identified V30 [7] [17], V40 [8] [18], V50 [10] [19] [22], V55 [19], maximum heart dose [18], or mean heart dose [32]. The variability among studies may be attributable to strong correlations between DVH parameters which are caused by the physical properties of radiation beams (**Supplement S2.3**). The strength and pattern of such correlations may depend on treatment delivery techniques, which change over time and may thus affect each study differently. Advanced statistical techniques, such as the techniques employed in our study, confer an advantage of systematically examining the entire DVH, while accounting for correlations between DVH variables. Results in **Table 2** show that all  $V_D$  indices in the “sensitivity range” contribute to the Hazard Ratio. Additional discussion of potential reasons for discrepancies among studies is provided in the supplemental section (**Supplement S2.3**).

The etiology bridging the gap between heart dose and survival has yet to be confirmed. A recent systematic review details the existing literature on dosimetry and cardiac endpoints across pediatric, breast, lung, esophageal, and hematologic malignancies, emphasizing risk for coronary disease, valvular disease, arrhythmia, and pericardial disease after thoracic RT [24]. The apparent importance of upper heart substructures and specifically the superior right heart [9] [20] [30] [33] suggest radiation damage to the cardiac conduction system may impact survival in NSCLC patients. Several recent findings support this hypothesis. Among NSCLC patients treated on prospective dose-escalation trials at University of North Carolina, 11% had documented arrhythmia at 26 months after RT [3]. However, if radiation caused transient fatal arrhythmias, they would not likely be identified and may simply be recorded as deaths due to lung cancer. At 6 months after thoracic RT for locally advanced NSCLC, Vivekanandan *et al.* [9] found ECG changes in 38% of patients and ECG changes were associated with worse OS on multivariate analysis. Adding to this evidence, we have recently published [33] an expanded analysis of 3-dimensional dose distributions in the heart, for the same patient cohort as the present study, which found that the dose to the right-superior portion of the heart was most responsible for the decreased OS. More detailed cardiac evaluation of NSCLC patients receiving thoracic RT is necessary to evaluate this hypothesis.

We acknowledge the limitations of our study. Although our findings were statistically significant, our sample size is limited. Of the 16 studies in the past 10 years that have found an association between heart dose and survival, 12 included more patients than the present study (**Table 5**) [4] [7] [8] [10] [17]-[23] [28] [29] [30] [31]. Moreover, our data only included the clinical outcome of OS without any other clinical outcomes like cardiac events. Multivariate analysis mitigates the possibility of confounding by other disease and treatment-related variables but cannot exclude confounding by unaccounted for variables. Some studies have suggested confounding by or interactions with immunosuppression [22], pre-existing coronary heart disease [32], lung dose [34], or extent of mediastinal lymph node involvement [35]. Because of limited sample size we only

performed Leave One Out Cross Validation and were not able to perform sample subdivision into model fitting and validation parts. Additional validation must be left to future work with an expanded data sample.

Based on our findings and the existing literature, high dose to the heart should be avoided whenever possible. For patients with NSCLC treated with conventionally fractionated RT, heart doses  $\geq 51$  Gy may decrease OS. The superior right heart may be the most at-risk for radiation induced toxicity. Data shown in **Table 2** and **Table 4** (heart DVH, age, stage, receipt of chemotherapy) can be used to calculate an individualized HR for OS for every stage III NSCLC patient undergoing thoracic RT (**Supplement S1**).

## 5. Conclusion

Among stage III NSCLC patients undergoing thoracic RT, worse OS is associated with higher heart dose, older age, more advanced stage, and lack of chemotherapy. Doses higher than 51 Gy were predictive for reduced OS, while heart V55 appeared to provide the best estimate of OS for setting treatment planning constraints, with HR 1.044 per 1% increase of heart volume exposed to at least 55 Gy.

## Acknowledgements

Mayo Clinic Arizona and the Arizona State University Rising Star Initiative provided funding to support the collaboration between the Department of Radiation Oncology at Mayo Clinic Arizona and the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University. Varian, Inc. provided a grant to Arizona State University to fund this work. SES received support for research time from the North Central Cancer Treatment Group and Mayo Clinic with funding from the Public Health Service (CA-25224, CA-37404, CA-35267, CA-35431, CA-35195, CA-63848, CA-63849, CA-35113, CA-35103, CA-35415, CA-35101, CA-35119, CA-35090).

## Statistical Analysis

Prof. Jing Li supervised the statistical analysis; Jiuyun Hu carried out the analysis as part of his PhD dissertation.

## Data Sharing

Research data not available at this time.

## Conflicts of Interest

Dr. Schild reports writing and editing for UpToDate. No conflicts exist for remaining authors.

## References

- [1] Lancellotti, P., Nkomo, V.T., Badano, L.P., Bergler-Klein, J., Bogaert, J., Davin, L., *et*

- al.* (2013) Expert Consensus for Multi-Modality Imaging Evaluation of Cardiovascular Complications of Radiotherapy in Adults: A Report from the European Association of Cardiovascular Imaging and the American Society of Echocardiography. *European Heart Journal: Cardiovascular Imaging*, **14**, 721-740. <https://doi.org/10.1093/ehjci/jet123>
- [2] Dess, R.T., Sun, Y., Matuszak, M.M., Sun, G., Soni, P.D., Bazzi, L., *et al.* (2017) Cardiac Events after Radiation Therapy: Combined Analysis of Prospective Multicenter Trials for Locally Advanced Non-Small-Cell Lung Cancer. *Journal of Clinical Oncology*, **35**, 1395-1402. <https://doi.org/10.1200/JCO.2016.71.6142>
- [3] Wang, K., Eblan, M.J., Deal, A.M., Lipner, M., Zagar, T.M., Wang, Y., *et al.* (2017) Cardiac Toxicity after Radiotherapy for Stage III Non-Small-Cell Lung Cancer: Pooled Analysis of Dose-Escalation Trials Delivering 70 to 90 Gy. *Journal of Clinical Oncology*, **35**, 1387-1394. <https://doi.org/10.1200/JCO.2016.70.0229>
- [4] Xue, J., Han, C., Jackson, A., Hu, C., Yao, H., Wang, W., *et al.* (2019) Doses of Radiation to the Pericardium, Instead of Heart, are Significant for Survival in Patients with Non-Small Cell Lung Cancer. *Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology*, **133**, 213-219. <https://doi.org/10.1016/j.radonc.2018.10.029>
- [5] Wang, K., Pearlstein, K.A., Patchett, N.D., Deal, A.M., Mavroidis, P., Jensen, B.C., *et al.* (2017) Heart Dosimetric Analysis of Three Types of Cardiac Toxicity in Patients Treated on Dose-Escalation Trials for Stage III Non-Small-Cell Lung Cancer. *Radiotherapy and Oncology*, **125**, 293-300. <https://doi.org/10.1016/j.radonc.2017.10.001>
- [6] Ning, M.S., Tang, L., Gomez, D.R., Xu, T., Luo, Y., Huo, J., *et al.* (2017) Incidence and Predictors of Pericardial Effusion After Chemoradiation Therapy for Locally Advanced Non-Small Cell Lung Cancer. *International Journal of Radiation Oncology, Biology, Physics*, **99**, 70-79. <https://doi.org/10.1016/j.ijrobp.2017.05.022>
- [7] Bradley, J.D., Paulus, R., Komaki, R., Masters, G., Blumenschein, G., Schild, S., *et al.* (2015) 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 Phase 3 Study. *The Lancet Oncology*, **16**, 187-199. [https://doi.org/10.1016/S1470-2045\(14\)71207-0](https://doi.org/10.1016/S1470-2045(14)71207-0)
- [8] Chun, S.G., Hu, C., Choy, H., Komaki, R.U., Timmerman, R.D., Schild, S.E., *et al.* (2017) Impact of Intensity-Modulated Radiation Therapy Technique for Locally Advanced Non-Small-Cell Lung Cancer: A Secondary Analysis of the NRG Oncology RTOG 0617 Randomized Clinical Trial. *Journal of Clinical Oncology*, **35**, 56-62. <https://doi.org/10.1200/JCO.2016.69.1378>
- [9] Vivekanandan, S., Landau, D.B., Counsell, N., Warren, D.R., Khwanda, A., Rosen, S.D., *et al.* (2017) The Impact of Cardiac Radiation Dosimetry on Survival after Radiation Therapy for Non-Small Cell Lung Cancer. *International Journal of Radiation Oncology, Biology, Physics*, **99**, 51-60. <https://doi.org/10.1016/j.ijrobp.2017.04.026>
- [10] Eaton, B.R., Pugh, S.L., Bradley, J.D., Masters, G., Kavadi, V.S., Narayan, S., *et al.* (2016) Institutional Enrollment and Survival Among NSCLC Patients Receiving Chemoradiation: NRG Oncology Radiation Therapy Oncology Group (RTOG) 0617. *Journal of the National Cancer Institute*, **108**, Article ID: djw034. <https://doi.org/10.1093/jnci/djw034>
- [11] Zhang, T.W., Snir, J., Boldt, R., Rodrigues, G., Louie, A., Gaede, S., *et al.* (2019) Is the Importance of Heart Dose Overstated in the Treatment of Non-Small Cell Lung

- Cancer? A Systematic Review of the Literature. *International Journal of Radiation Oncology, Biology, Physics*, **104**, 582-589.  
<https://doi.org/10.1016/j.ijrobp.2018.12.044>
- [12] Chong, I. and Jun, C. (2005) Performance of Some Variable Selection Methods When Multicollinearity Is Present. *Chemometrics and Intelligent Laboratory Systems*, **78**, 103-112. <https://doi.org/10.1016/j.chemolab.2004.12.011>
- [13] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**, 267-288.  
<https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [14] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005) Sparsity and Smoothness via the Fused Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **67**, 91-108. <https://doi.org/10.1111/j.1467-9868.2005.00490.x>
- [15] Dai, B. and Breheny, P. (2019) Cross Validation for Penalized Cox Regression. arXiv: 1905. 10432v1.
- [16] Zou, H. and Hastie, T. (2015) Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society Series B (Methodological)*, **67**, 301-320.  
<https://doi.org/10.1111/j.1467-9868.2005.00503.x>
- [17] Johnson, M.D., Sura, K., Mangona, V.S., Glick, A., Wallace, M., Ye, H., *et al.* (2017) Matched-Pair Analysis of High Dose Versus Standard Dose Definitive Chemoradiation for Locally Advanced Non-Small-Cell Lung Cancer. *Clinical Lung Cancer*, **18**, 149-155. <https://doi.org/10.1016/j.clcc.2016.06.004>
- [18] Sio, T.T., Liang, J.J., Chang, K., Jayakrishnan, R., Novotny, P.J., Prasad, A., *et al.* (2017) Dosimetric Correlate of Cardiac-Specific Survival among Patients Undergoing Coronary Artery Stenting after Thoracic Radiotherapy for Cancer. *American Journal of Clinical Oncology: Cancer Clinical Trials*, **40**, 133-139.  
<https://doi.org/10.1097/COC.000000000000135>
- [19] Speirs, C.K., DeWees, T.A., Rehman, S., Molotievschi, A., Velez, M.A., Mullen, D., *et al.* (2017) Heart Dose Is an Independent Dosimetric Predictor of Overall Survival in Locally Advanced Non-Small Cell Lung Cancer. *Journal of Thoracic Oncology*, **12**, 293-301. <https://doi.org/10.1016/j.jtho.2016.09.134>
- [20] Stam, B., Peulen, H., Guckenberger, M., Mantel, F., Hope, A., Werner-Wasik, M., *et al.* (2017) Dose to Heart Substructures Is Associated with Non-Cancer Death after SBRT in Stage I-II NSCLC Patients. *Radiotherapy and Oncology*, **123**, 370-375.  
<https://doi.org/10.1016/j.radonc.2017.04.017>
- [21] Vivekanandan, S., Landau, D.B., Counsell, N., Warren, D.R., Khwanda, A., Rosen, S.D., *et al.* (2017) The Impact of Cardiac Radiation Dosimetry on Survival after Radiation Therapy for Non-Small Cell Lung Cancer. *International Journal of Radiation Oncology, Biology, Physics*, **99**, 51-60.  
<https://doi.org/10.1016/j.ijrobp.2017.04.026>
- [22] Contreras, J.A., Lin, A.J., Weiner, A., Speirs, C., Samson, P., Mullen, D., *et al.* (2018) Cardiac Dose is Associated with Immunosuppression and Poor Survival in Locally Advanced Non-Small Cell Lung Cancer. *Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology*, **128**, 498-504.  
<https://doi.org/10.1016/j.radonc.2018.05.017>
- [23] Wong, O.Y., Yau, V., Kang, J., Glick, D., Lindsay, P., Le, L.W., *et al.* (2018) Survival Impact of Cardiac Dose Following Lung Stereotactic Body Radiotherapy. *Clinical Lung Cancer*, **19**, E241-E246. <https://doi.org/10.1016/j.clcc.2017.08.002>
- [24] Niska, J.R., Thorpe, C.S., Allen, S.M., Daniels, T.B., Rule, W.G., Schild, S.E., *et al.*

- (2018) Radiation and the Heart: Systematic Review of Dosimetry and Cardiac Endpoints. *Expert Review of Cardiovascular Therapy*, **16**, 931-950. <https://doi.org/10.1080/14779072.2018.1538785>
- [25] Ramsey, S.D., Howlader, N., Etzioni, R.D. and Donato, B. (2004) Chemotherapy Use, Outcomes, and Costs for Older Persons with Advanced Non-Small-Cell Lung Cancer: Evidence from Surveillance, Epidemiology and End Results-Medicare. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **22**, 4971-4978. <https://doi.org/10.1200/JCO.2004.05.031>
- [26] Rami-Porta, R., Asamura, H., Travis, W.D. and Rusch, V.W. (2017) Lung Cancer—Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA: A Cancer Journal for Clinicians*, **67**, 138-155. <https://doi.org/10.3322/caac.21390>
- [27] Auperin, A., Le Pechoux, C., Pignon, J.P., Koning, C., Jeremic, B., Clamon, G., et al. (2006) Concomitant Radio-Chemotherapy Based on Platin Compounds in Patients with Locally Advanced Non-Small Cell Lung Cancer (NSCLC): A Meta-Analysis of Individual Data from 1764 Patients. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, **17**, 473-483. <https://doi.org/10.1093/annonc/mdj117>
- [28] Schytte, T., Hansen, O., Stolberg-Rohr, T. and Brink, C. (2010) Cardiac Toxicity and Radiation Dose to the Heart in Definitive Treated Non-Small Cell Lung Cancer. *Acta Oncologica*, **49**, 1058-1060. <https://doi.org/10.3109/0284186X.2010.504736>
- [29] Tukenova, M., Guibout, C., Oberlin, O., Doyon, F., Mousannif, A., Haddy, N., et al. (2010) Role of Cancer Treatment in Long-Term overall and Cardiovascular Mortality after Childhood Cancer. *Journal of Clinical Oncology*, **28**, 1308-1315. <https://doi.org/10.1200/JCO.2008.20.2267>
- [30] McWilliam, A., Kennedy, J., Hodgson, C., Vasquez, Osorio, E., Faivre-Finn, C. and van Herk, M. (2017) Radiation Dose to Heart Base Linked with Poorer Survival in Lung Cancer Patients. *European Journal of Cancer*, **85**, 106-113. <https://doi.org/10.1016/j.ejca.2017.07.053>
- [31] Taylor, C., Correa, C., Duane, F.K., Aznar, M.C., Anderson, S.J., Bergh, J., et al. (2017) Estimating the Risks of Breast Cancer Radiotherapy: Evidence from Modern Radiation Doses to the Lungs and Heart and from Previous Randomized Trials. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, **35**, 1641-1649. <https://doi.org/10.1200/JCO.2016.72.0722>
- [32] Atkins, K.M., Rawal, B., Chaunzwa, T.L., Lamba, N., Bitterman, D.S., Williams, C.L., et al. (2019) Cardiac Radiation Dose, Cardiac Disease, and Mortality in Patients with Lung Cancer. *Journal of the American College of Cardiology*, **73**, 2976-2987. <https://doi.org/10.1016/j.jacc.2019.03.500>
- [33] Liu, X., Fatyga, M., Schild, S.E. and Li, J. (2020) Detecting Spatial Susceptibility to Cardiac Toxicity of Radiation Therapy for Lung Cancer. *IJSE Transactions on Healthcare Systems Engineering*, **10**, 243-250. <https://doi.org/10.1080/24725579.2020.1795012>
- [34] Tucker, S.L., Liu, A., Gomez, D., Tang, L.L., Allen, P., Yang, J., et al. (2016) Impact of Heart and Lung Dose on Early Survival in Patients with Non-Small Cell Lung Cancer Treated with Chemoradiation. *Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology*, **119**, 495-500. <https://doi.org/10.1016/j.radonc.2016.04.025>
- [35] McNew, L.K., Bowen, S.R., Gopan, O., Nyflot, M.J., Patel, S.A., Zeng, J., et al. (2017)

The Relationship between Cardiac Radiation Dose and Mediastinal Lymph Node Involvement in Stage III Non-Small Cell Lung Cancer Patients. *Advances in Radiation Oncology*, **2**, 192-196. <https://doi.org/10.1016/j.adro.2017.01.008>

## Supplemental

### S1. Clinical Applications Supplement

The Cox proportional hazards regression model can be written as follows:

$$h(t) = h_0(t) \exp(b_1 X_1 + b_2 X_2 + \dots + b_n X_n)$$

where  $b_i X_i$  represent potential predictors and  $h_0(t)$  a baseline risk when all predictors are set to zero.

#### S1.1. Whole Heart DVH Analysis

**Table S1** shows fit coefficients, hazard ratios, and p-values which were associated with significant predictors of overall survival in the model which included V55Gy for the whole heart DVH.

#### S1.2. Example Applications Using Whole Heart V55Gy

1) With the “reference patient” defined as stage IIIA, no chemotherapy, 70.5 years old, and heart V55Gy = 0%, the individualized hazard ratio for an actual patient with stage IIIB disease, who received chemotherapy, and whose heart V55Gy was  $N\%$ , can be computed as (from **Table 4**):

$$HR = 1.78 \times 0.46 \times 1.04^{(\text{age}-70.5)} \times 1.044^N$$

The interpretation of  $HR = X$  is that the likelihood of dying per unit of time for the actual patient is  $X$ -fold greater than the likelihood of dying per unit of time for the reference patient.

2) Holding the other variables constant, for a patient whose V55Gy to the whole heart is  $N\%$ , the associated hazard ratio for cardiac radiation exposure can be calculated as

$$e^{N \times 0.043} \text{ or } 1.044^N.$$

Consider two patients who are the same age, have the same cancer stage, and both received chemotherapy, but differ only by extent of heart irradiation. Patient A received heart V55Gy = 0%. Patient B received heart V55Gy = 10%. Using the coefficient from the second column of **Table S1**, the individualized hazard ratio for patient A is

$e^{0 \times 0.043} = 1.0$ . The individualized hazard ratio for patient B is  $e^{10 \times 0.043} = 1.537$ . The likelihood of dying per unit of time is 1.537 times higher for patient B than for patient A.

**Table S1.** Model coefficients and associated hazard ratios in the model with whole heart DVH.

Predictor $X_i$	Coefficient $b_i$	Hazard Ratio	p-value
Stage IIIB	0.57	1.78	0.016
Chemotherapy Use	-0.79	0.46	0.039
Age before RT	0.04	1.04/1 year	0.012
V55Gy	0.0430	1.044/1% of volume	0.034

Another way to obtain the estimate of hazard ratio associated with cardiac irradiation is to use the hazard ratio listed in **Table 4**. Using the same example patients as before, the hazard ratio for patient B is  $1.044^{10} = 1.538$ .

3) If two groups of patients have hazard ratios equal to  $HR_1$  and  $HR_2$ , the ratio of the mean survival times between the two groups can be approximated by  $\frac{HR_1}{HR_2}$ .

## S2. Statistical Analysis Supplement

### S2.1. Radiobiological and Clinical Motivation for “Knowledge Constraints” in KC-Lasso

#### 1) Introduction

Dose-volume analysis is one of the primary tools used in phenomenological modelling of clinical toxicity in radiation therapy. Dose volume analysis reflects the basic clinical and radiobiological insight that the likelihood of clinical toxicity depends on both the dose level and the volume to which the dose is applied. In general, larger doses and larger volumes to which dose is applied lead to greater likelihood of toxicity. The dominant effect of depositing dose in a volume of tissue is cell kill (cell depletion), usually described as cell Survival Fraction (SF), which is the proportion of cells that survive an irradiation. The SF is related to dose by a Linear Quadratic model equation  $SF = e^{-\alpha \cdot BED}$  where BED is Biologically Equivalent Dose (BED) related to the physical dose through a linear quadratic equation  $BED = D \left( 1 + \frac{d}{\alpha/\beta} \right)$ . In the discussion that follows we will equate the physical dose  $D$  with BED and refer to both as “dose”.

Irradiating a volume of tissue with a dose level “ $D$ ” can lead to one of the three categories of outcomes:

- a) The SF may be high enough that the tissue can compensate for lost cells and there is no clinical toxicity observed.
- b) The SF is low enough that the tissue may not be able to fully compensate for lost cells which can lead to transient or permanent toxicity. The toxicity is transient if the tissue can rebuild itself over time and permanent if the tissue can no longer rebuild itself. In this regime, the onset of toxicity is probabilistic and may additionally depend on patient specific characteristics, like age, state of health or genetics.
- c) The SF is so low that toxicity will inevitably occur, for all patients.

The (1)-(3) states are typically modelled by the sigmoid (logistic) curve. The Region (1) corresponds to the beginning part of the curve, Region (2) corresponds to the rising part of the curve, and Region (3) corresponds to the saturation region of the logistic curve.

#### 2) The Linear Predictor in KC-Lasso

The linear predictor in KC-Lasso assumes the following form

$$\eta = \beta_0 + \beta_1 V_{D_1} + \dots + \beta_p V_{D_p}$$

where  $V_{D_i}$  is the percentage of organ volume with dose  $D_i$ , or greater.

### 3) Positivity Condition in KC-Lasso

The positivity condition in KC-Lasso says that all coefficients  $\beta_i$  have to be greater or equal to zero. ( $\beta \geq 0$ )

The motivation for the positivity condition is that any dose, applied to any volume, kills cells ( $SF < 1$ ). Such irradiation can produce no risk (Region (1)), some risk (Regions (2)-(3)), but it cannot reduce risk, which would be implied by negative coefficients.

### 4) Monotonicity Condition in KC-Lasso

Consider one of the elements in the linear predictor:

$$\beta_i V_{D_i}$$

Let us fix the value of  $V_{D_i}$  at an arbitrarily chosen value of 0.05 (5% of the volume irradiated to the dose  $D_i$  or higher). What should the contribution of this term to risk be when  $D_i$  increases? An increase in dose means that the surviving fraction (SF) decreases. A decrease in SF means that the contribution of this term to risk has to increase (Region (2), increasing risk) or stay the same (Region (3), saturated risk).

Returning to the full linear predictor feature, one thus observes that, for a fixed value of  $V_{D_i}$  (e.g. 0.05) the contribution of consecutive terms to the feature has to increase, or stay the same, as the dose increases. This argument leads us to the monotonicity condition, which says that consecutive coefficients must be greater or the same as their predecessors:

For any two dose levels  $d_1, d_2$ , with  $d_1 \leq d_2$ , let  $\beta_{d_1}$  and  $\beta_{d_2}$  be the coefficients for  $V_{d_1}$  and  $V_{d_2}$ , respectively. Then,  $\beta_{d_1} \leq \beta_{d_2}$ .

## S2.2. Internal Validation

Since the KC lasso method is based on fused lasso, there are two tuning parameters we need to determine, *i.e.*, the L1 penalty for the coefficients and the L1 penalty for the steps of the coefficients. We use the work by Dai and Breheny [15] who concluded that a method of leave-one-out cross validation with linear predictors works the best for a Cox model. To be more specific, suppose that there are  $n$  patients in the data. In the  $i$ -th iteration, we leave the  $i$ -th patient out and use the remaining  $n-1$  patients to fit the Cox model with KC lasso. Then the ‘linear predictor’ of the  $i$ -th patient can be defined as  $\hat{\eta}_i = \bar{X}_i^T \hat{\beta}^{-i}$ , where  $\bar{X}_i^T$  is the feature vector for  $i$ -th patient and  $\hat{\beta}^{-i}$  is the estimated coefficient vector of the model without  $i$ -th patient. After we get “linear predictors” for all patients, we can define the predictive accuracy based on the “linear predictors” by

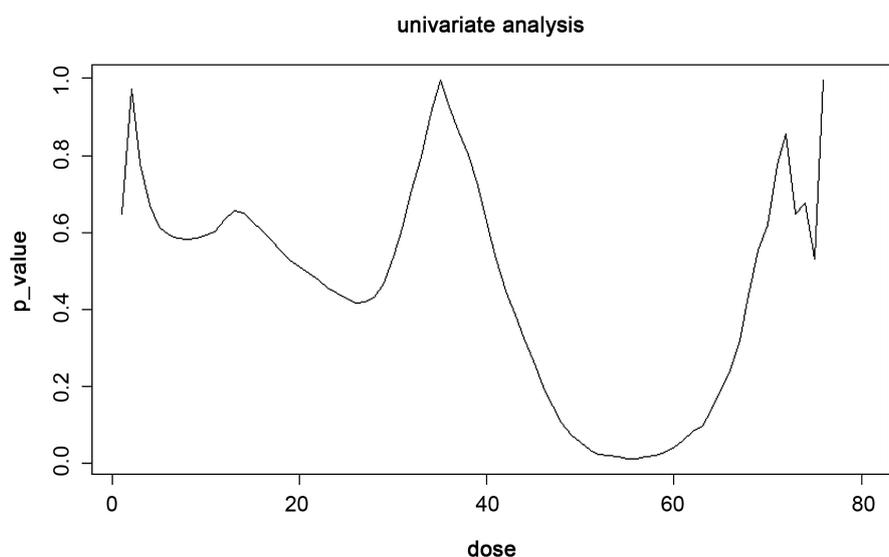
$$L(\hat{\eta}) = \prod_{i=1}^n \left[ e^{\hat{\eta}_i} / \sum_{j \in R(t_i)} e^{\hat{\eta}_j} \right],$$

where  $t_i$  is the event time or last follow up time of  $i$ -th patient.  $R(t_i)$  is the set of patients at risk at time  $t_i$ . The cross validation error for linear predictors is then defined by  $CVE = -2 \log(L(\hat{\eta}))$ . We find the set of tuning parameters with the lowest CVE (cross validation error).

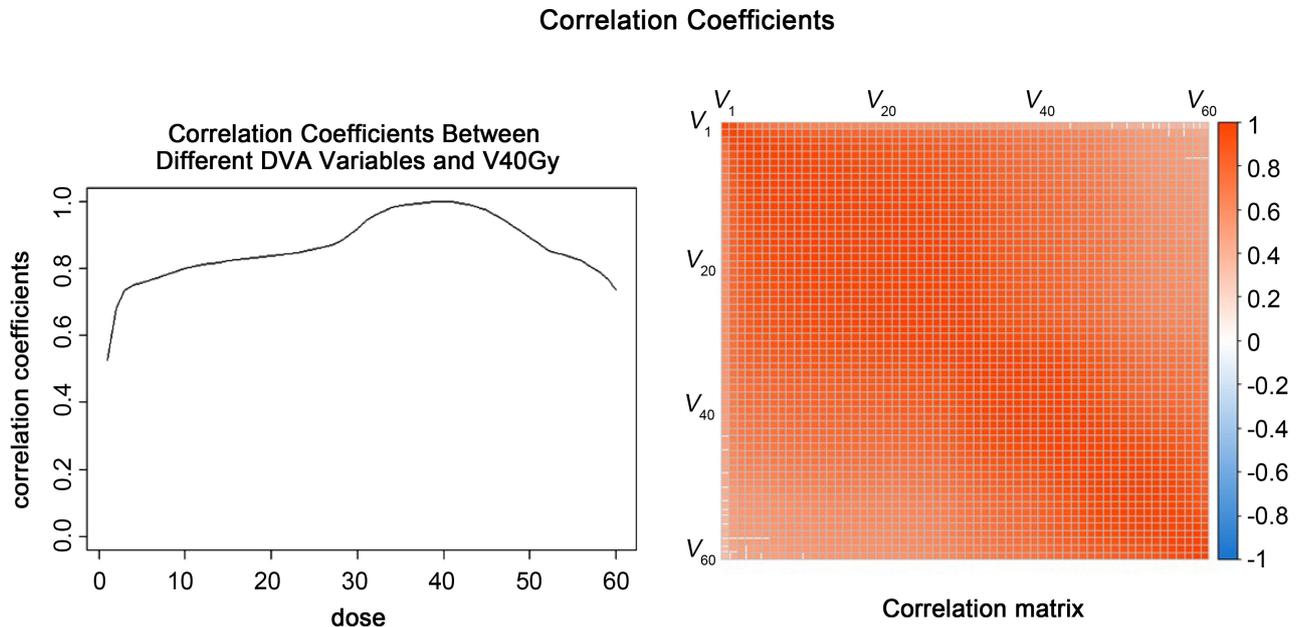
### S2.3. Limitations of the Univariate Analysis

In **Figure S1**, we show p-values associated with the dosimetric variable  $V_D$  in a Cox model which contains patient specific covariates (age, chemotherapy and stage) and one  $V_D$  index at a time. We effectively scanned the  $V_D$  space and recorded the p-value associated with each  $V_D$ . One observes a monotonic decline of p-values, starting between doses of 35 Gy - 40 Gy and continuing towards the minimum near  $V_{55\text{Gy}}$ . The monotonic decline in such a broad range of doses is caused primarily by correlations among indices for this population of patients, convolved with the threshold dose. The magnitude of correlations in the present data is shown in **Figure S2**. Correlations among indices are determined by the combination of three factors: 1) anatomic variability among patients, 2) variability in tumor location and volume, 3) dose distributions which are characteristic for the delivery methods being used. The magnitude of p-value at the minimum also depends on the sample size. If one sets an arbitrary threshold of p-value at  $p = 0.05$ , this threshold will be crossed at a dose level which depends both on the sample size and on the pattern of correlations, which can vary in different studies due to differences in treatment delivery methods being used. Hence, the method of searching for the dose threshold with univariate analysis tends to produce threshold doses which are too low and can vary among studies. A more advanced statistical technique, which explicitly “corrects” for correlations, needs to be used to detect a threshold dose which will remain consistent for all studies and is generalizable to future clinical applications.

If one considers multiple  $V_D$  indices as uncorrelated, independent variables, p-value obtained for each of these variables should be subjected to multiple comparisons correction. The correction will depend on the size of the scanning



**Figure S1.** P-values associated with  $V_D$  covariate in a Cox model with patient specific variables (age, stage and chemotherapy) and one dosimetric covariate at a time. Doses are in Gy and a 1 Gy step was used in the analysis. Data in **Table 3** were sampled from this figure.



**Figure S2.** Correlations among  $V_D$  indices, the correlation matrix on the right side and a single section through the matrix at the level of  $V_{40}$  on the left side.

step and may prevent the p-values from reaching the threshold of statistical significance for patient cohorts of a realistic size. Adjusting the step size to reach the threshold of statistical significance is not well justified and reduces the precision of searching for a dose threshold which is associated with the clinical outcome. The magnitude of this problem is illustrated in **Figure S3**, using the  $V_D$  step size of 1 Gy. Red bars show p-values obtained without multiple comparison correction, while blue bars show p-values adjusted for the correction. The threshold of statistical significance would not have been reached in the present patient cohort if p-values were corrected for multiple comparisons.

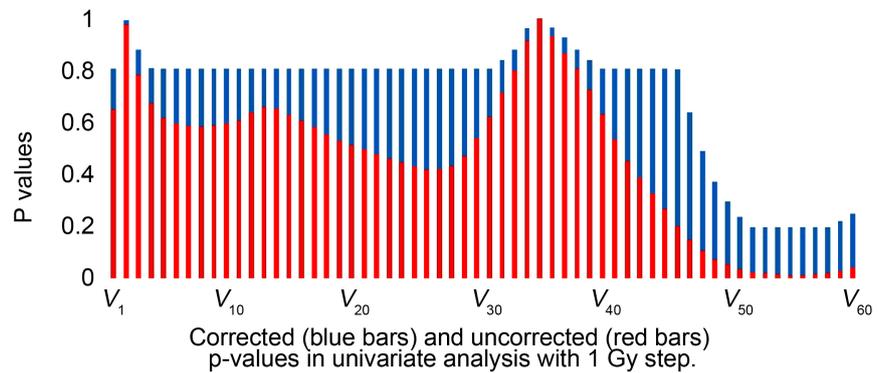
#### S2.4. Alternative Statistical Methods

KC-Lasso has been designed for the present study to address the problem of correlations among dosimetric variables. One can reasonably ask whether other statistical techniques could perform equally as well as KC-Lasso model. To address this question we compared KC-Lasso (Knowledge Constrained, Fused Lasso) to Elastic Net [16], Lasso [13] and Fused Lasso [14] models. We make a comparison by first deriving the coefficients for the  $V_D$  variables in each model which creates a linear predictor (feature) for each. We then compare the p-value associated with the feature in an unpenalized Cox model containing patient specific covariates. All four models were associated with very similar p-values, as summarized in **Table S2**.

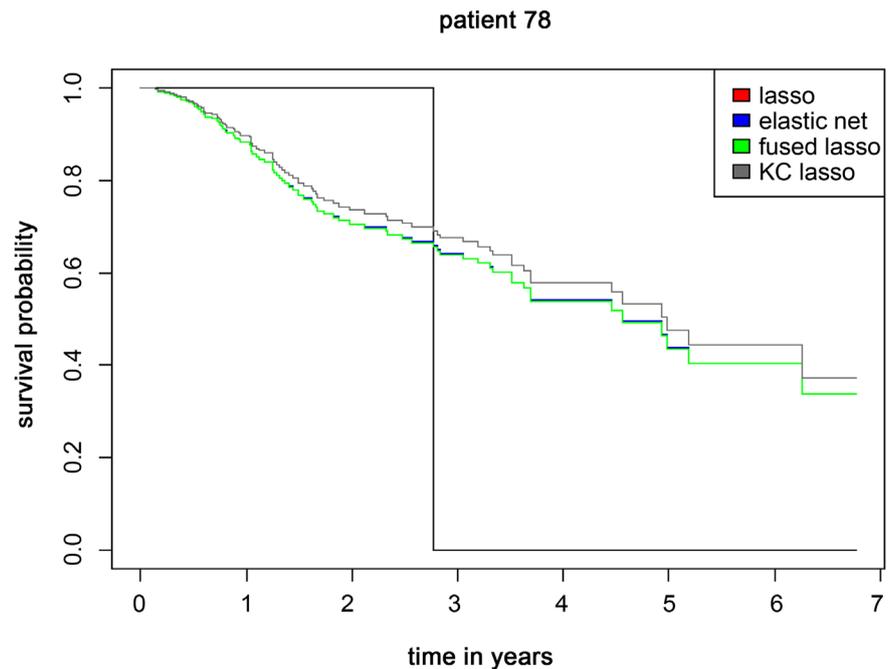
All four models produced very similar survival curves for surviving patients. KC-Lasso showed some differences in survival curves for deceased patients. Two examples are shown in **Figure S4** and **Figure S5**.

**Table S2.** P-values associated with the linear predictor in four models in an unpenalized cox model with patient specific covariates.

Model	p-value
KC-Lasso	0.0122
Lasso	0.0134
Fused Lasso	0.0123
Elastic Net	0.0122

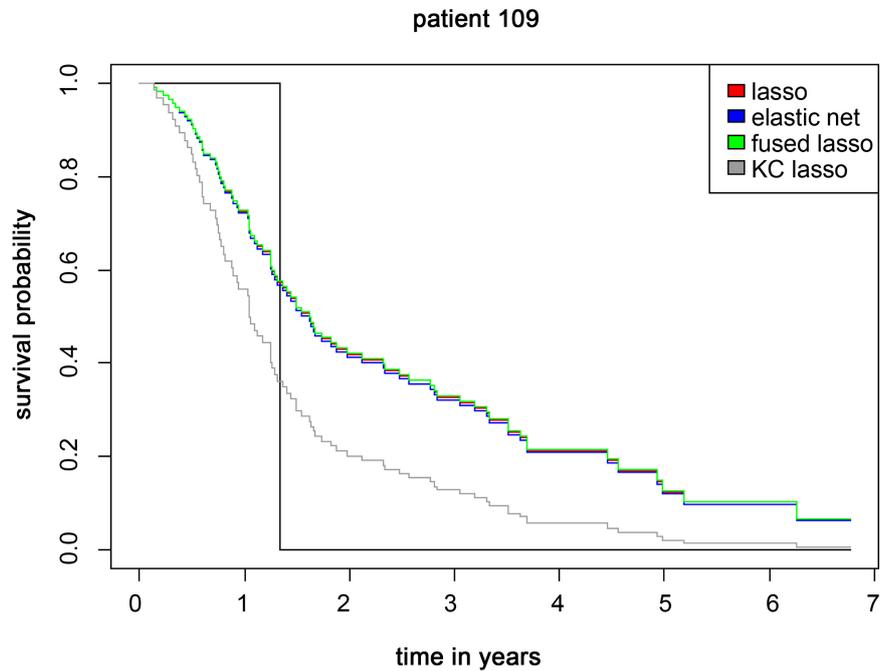


**Figure S3.** An illustration of the effect that multiple comparisons correction would have on p-values in the univariate analysis for the present study.



**Figure S4.** An example of survival curves for a surviving patient.

All four models perform similarly on the same data set, though one could argue that KC-Lasso is showing slightly better prediction of survival probability. The primary difference between the models is in the choice of coefficients for the linear predictors that each model makes. The coefficients chosen by the Elastic



**Figure S5.** An example of survival curves for a deceased patient.

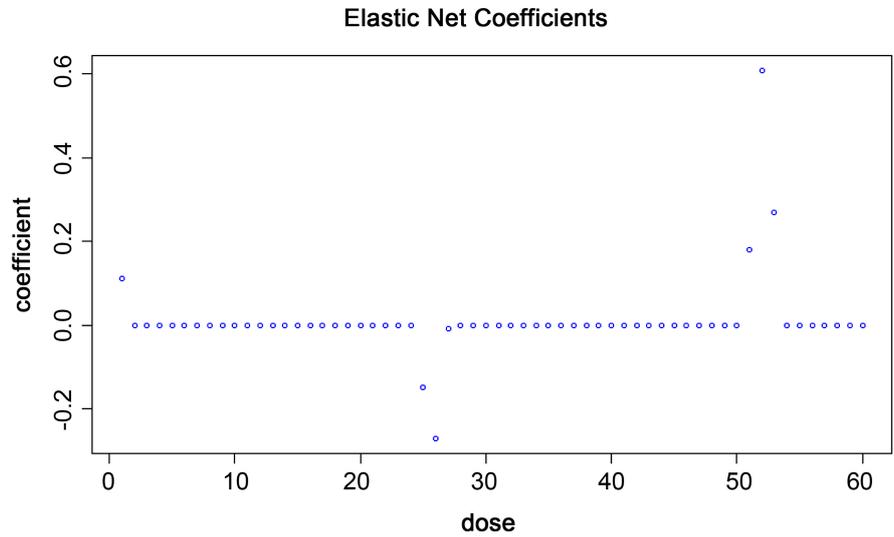
Net (**Figure S6**), Lasso (**Figure S7**) and Fused Lasso (**Figure S8**) models are shown below.

One observes that both the Elastic Net, Lasso and Fused Lasso models selected negative coefficients. Negative coefficients are biologically implausible because they imply a possibility that the irradiation of tissue at risk improves the odds of survival. A second biologically implausible feature is the selection of a single  $V_D$  index (or a small group of indices) without a simultaneous selection of indices at higher doses. Higher doses are always associated with lower cell survival fraction in the volume. Radiobiology suggests that lower cell survival fraction in a fixed volume of tissue should always be associated with higher likelihood of clinical complications, or at least the same likelihood of complications if cell survival fractions are so low that the adverse clinical outcome is virtually assured. Consequently, once the first  $V_{D_{thr}}$  (threshold dose) is selected, all  $V_{D > D_{thr}}$  should also be selected and their coefficients should be higher (more risk) or equal (saturated risk). The constraint on the maximum dose in the analysis must be imposed by the maximum dose available in the data.

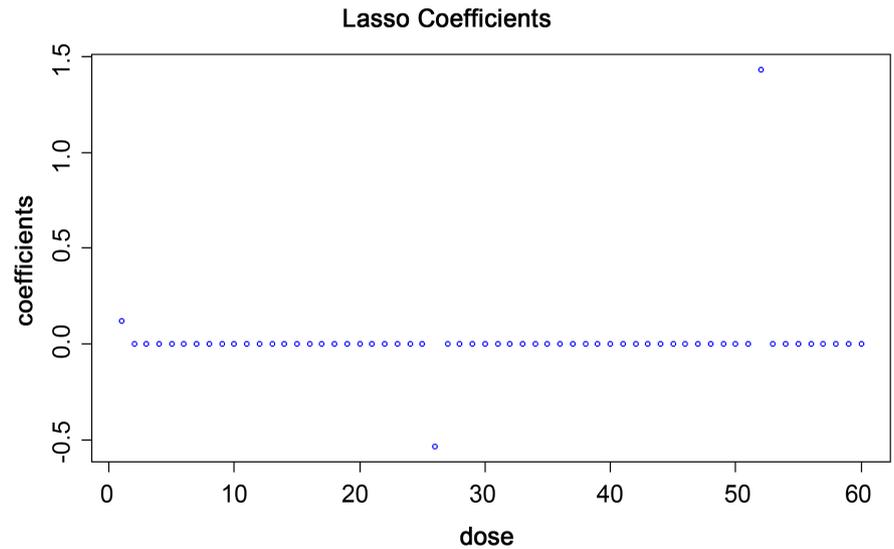
In summary, KC-Lasso is not the only statistical model that can be successfully fit to the present data set. However, KC-Lasso has been designed to satisfy “common sense” boundary conditions (positivity and monotonicity conditions imposed on coefficients) as well as to account for correlations between  $V_D$  indices. The purpose of this design has been to make the results of the model easier to interpret intuitively and to be more generalizable.

### S2.5. Tumor Volume as a Patient Specific Covariate

Tumor volume can influence the likelihood of OS and can also influence



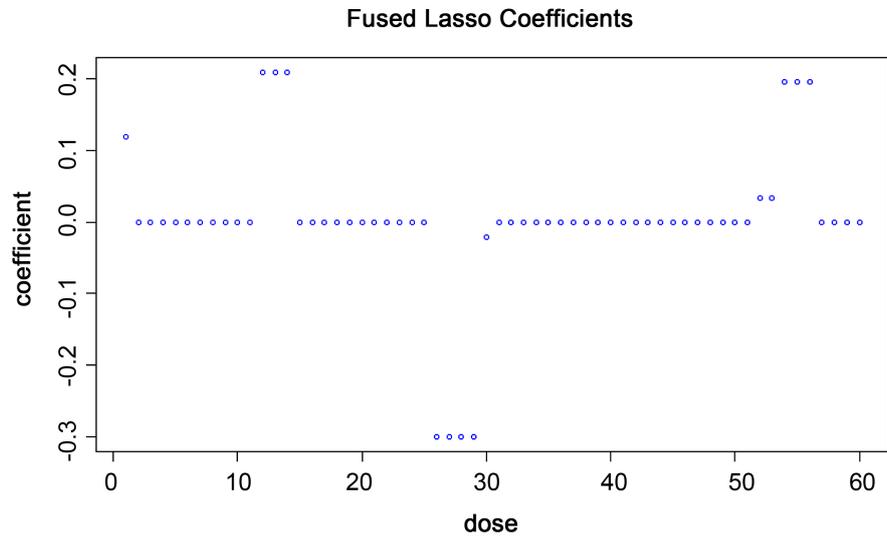
**Figure S6.** Coefficients selected by the elastic net model.



**Figure S7.** Coefficients selected by the lasso model.

the irradiation of the heart. Tumor volume was difficult to assess in this retrospective study because physicians frequently broke the target hierarchy (GTV-CTV-PTV-ITV) during contouring and thus forced estimates of the tumor volume with assumptions. We estimated the tumor volume (under assumptions) but decided to substitute tumor volume with a patient specific covariate which is strongly correlated with volume, namely clinical stage (3A and 3B). The clinical stage has been reliably recorded prior to treatment and reflects added clinical risk associated with tumor progression. The correlation between our estimated tumor volume and the clinical stage is summarized in **Table S3**.

When volume alone is used in the Cox model it is not a predictor for the OS ( $p = 0.35$ ). When volume alone is used in the Cox model with chemotherapy and age it is predictive for OS with  $p = 0.048$ . When stage is used instead of volume,



**Figure S8.** Coefficients selected by the fused lasso model.

**Table S3.** Summary of correlations between the disease stage and the estimated CTV volume.

Stage	Estimated CTV volume (cubic centimeters)			
	Mean	STDEV	MIN	MAX
3A	118.5	123.1	1.1	619
3B	180	143.6	4.7	706
t-test p-value	0.0083	N/A	N/A	N/A

the stage (3A/3B) is a predictor for OS with  $p = 0.024$ . When volume and stage are included simultaneously, neither one is the predictor with  $p$ -values  $p = 0.18$  for volume and  $p = 0.07$  for stage. Given that stage and estimated volume are strongly correlated, we chose to include stage alone in the analysis.