Department of Radiation Oncology

University of Virginia

Classification of Patient Risk from Stereotactic Body Radiation Therapy treating thoracic cancers

Author:

John Fishbein

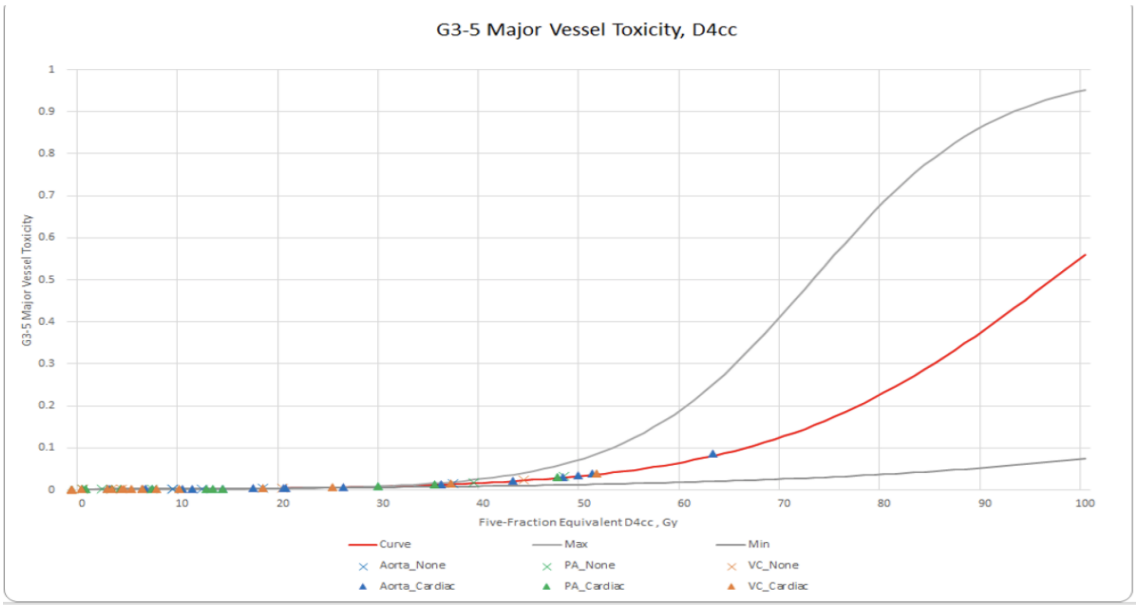
Supervisor:

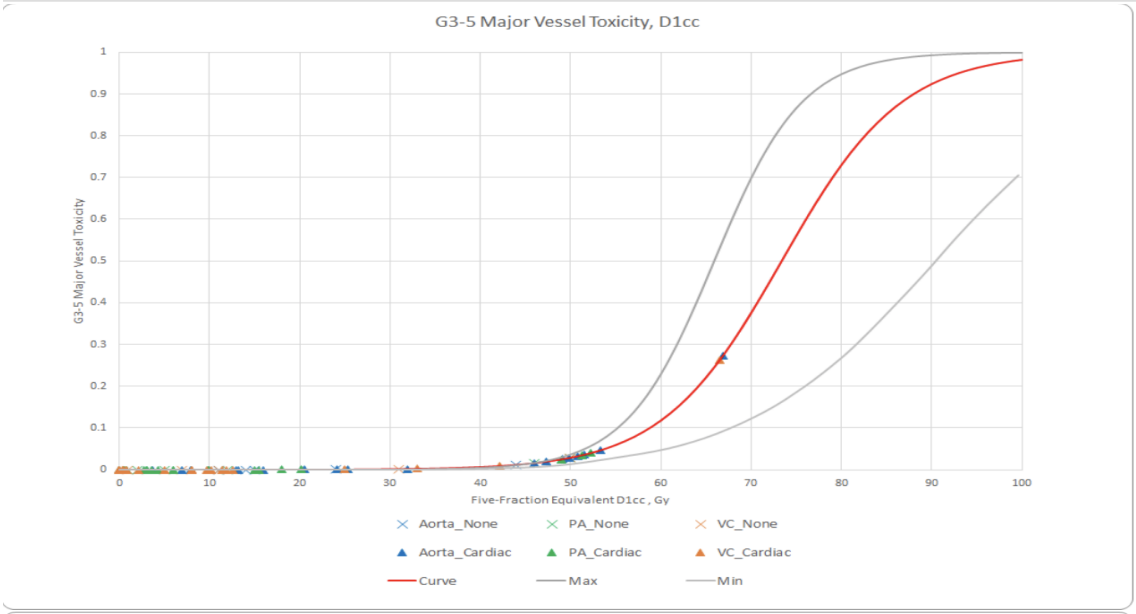
Dr. Krishni Wijesooriya

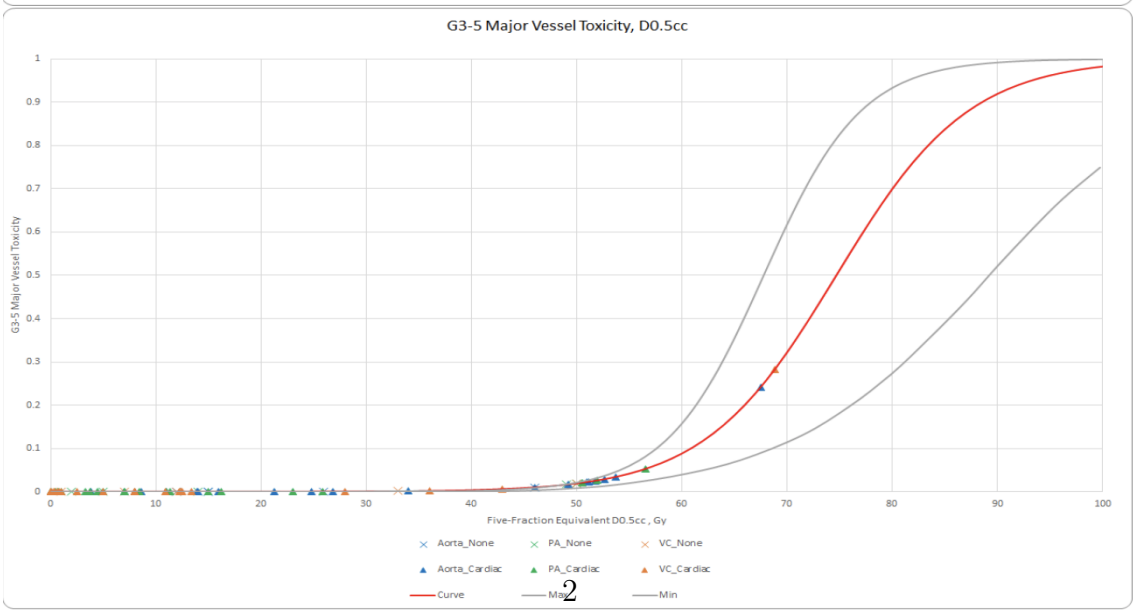
April 28, 2020

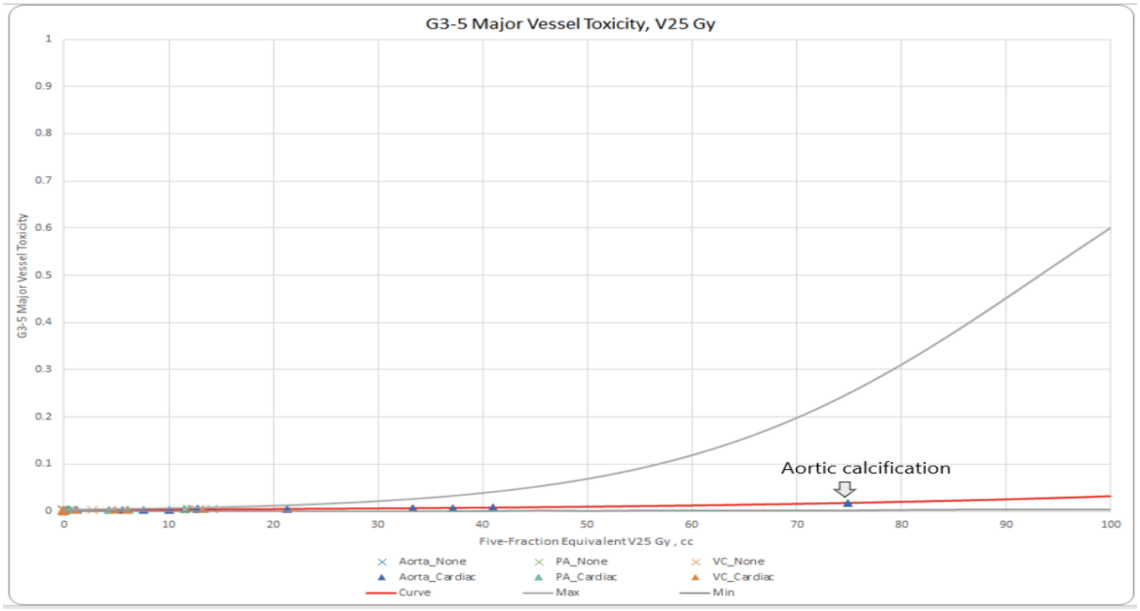
1 Introduction

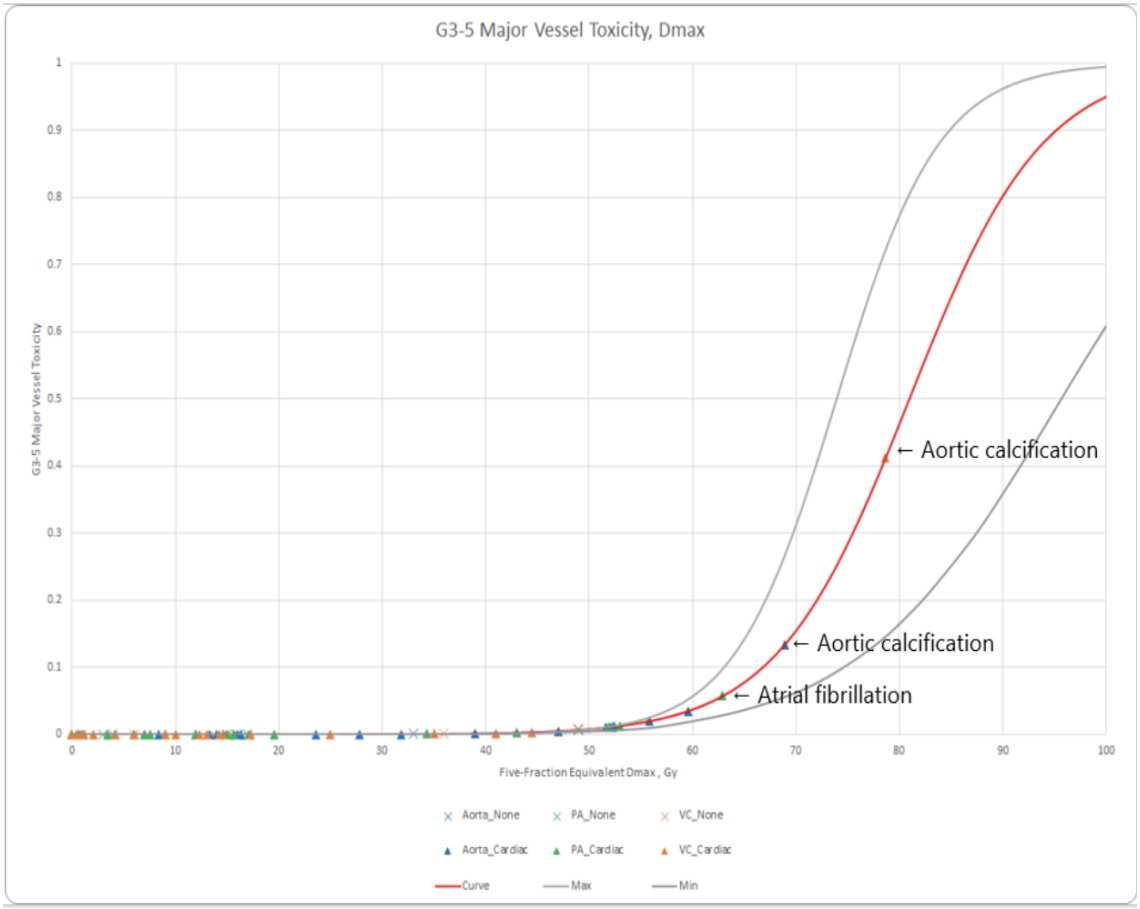
In practice today, current radiation oncologists treat early stage thoracic cancers using a technique known as Stereotactic Body Radiation Therapy (SBRT). SBRT works by delivering a massive dose of radiation, often in excess of 50Gy, to the Planning Target Volume (PTV) in 5 or fewer fractions and thus eradicating the tumor. This treatment is equivalent to surgery and has proven to have very high cure rates, but it has only been used by oncologists for the past 20 years. While definitive treatments with SBRT are curing many patients from their lung cancers, there are associated cardiovascular side effects of radiation that can impact patient's overall survival, quality of life, and healthcare costs. Indeed, lung cancer patients are already at risk for cardiovascular disease due to the high rates of smoking and environmental exposures in this patient population. Unfortunately, it is not possible to truly target only the tumors, and thus adjacent normal tissues inevitably receive some radiation. The dose and volume of radiation to these normal structures strongly impact the side-effects and depend largely on the location of the primary tumor. Development of a strategy to limit cardiac toxicity in SBRT is a complex issue. This is especially due to incomplete toxicity information with hypofractionated radiation, uncertainty in tolerance of radiation to different anatomical cardiac areas and adjacent major vessels, inter and intra-fraction cardiac motion, and patient specific radiation toxicity susceptibility factors. Currently, there are no models available to evaluate the overall cardiac toxicity due to a lung SBRT plan. As exemplified in the Xue et al. paper (1), there is a statistically significant correlation between these doses to cardiovascular structures under SBRT and further health complications. This must be considered when planning a safe and effective treatment.


G3-5 Major Vessel Toxicity, D4cc
G3-5 Major Vessel Toxicity
Five-Fraction Equivalent D4cc , Gy
Curve
Max
Min
Aorta_None
PA_None
VC_None
Aorta_Cardiac
PA_Cardiac
VC_Cardiac


G3-5 Major Vessel Toxicity, D1cc
G3-5 Major Vessel Toxicity
Five-Fraction Equivalent D1cc , Gy
Aorta_None
PA_None
VC_None
Aorta_Cardiac
PA_Cardiac
VC_Cardiac
Curve
Max
Min


G3-5 Major Vessel Toxicity, D0.5cc
G3-5 Major Vessel Toxicity
Five-Fraction Equivalent D0.5cc , Gy
Aorta_None
PA_None
VC_None
Aorta_Cardiac
PA_Cardiac
VC_Cardiac
Curve
Min


G3-5 Major Vessel Toxicity, V25 Gy
G3-5 Major Vessel Toxicity
Five-Fraction Equivalent V25 Gy , cc
Aortic calcification
Aorta_None
PA_None
VC_None
Aorta_Cardiac
PA_Cardiac
VC_Cardiac
Curve
Max
Min


G3-5 Major Vessel Toxicity, Dmax
G3-5 Major Vessel Toxicity
Five-Fraction Equivalent Dmax , Gy
← Aortic calcification
← Aortic calcification
← Atrial fibrillation
Aorta_None
PA_None
VC_None
Aorta_Cardiac
PA_Cardiac
VC_Cardiac
Curve
Max
Min

The images pictured above illustrate major vessel toxicity as it relates to increasing dose constraints D0.5cc, D1cc, D4cc, V25Gy and Dmax. D0.5cc, D1cc, and D4cc represent the dose value in Gy to 0.5cc, 1cc, and 4cc of the major blood vessels in the body respectively. Dmax is the maximum point dose received by a voxel (3D pixel) of the great vessels during the treatment plan. The red line indicates the assumed exponential logistic model correlating Vessel Toxicity with dose. In the case of cardiovascular toxicity with SBRT treatments, researchers have looked at 5 dosimetric parameters alone, which are: V25Gy, D0.5cc, D1cc, D4cc, and Dmax for the organ of interest (Xue et al. (1)). A log logistic curve can be readily described by a two parameter function as shown in the equation below, with one parameter describing the dose at which 50% of patients exhibit complications, D50, and the second parameter, g, the normalized dose-response gradient (Bentzen et al. (2)). These curves of normal tissue complication probability can be generated as functions of any of the five dose parameters described above (i.e. Dv).

$$NTCP = \frac{e^{(4g_{50_V} * (\frac{D_V}{TD_{50_V}} - 1))}}{1 + e^{(4g_{50_V} (\frac{D_V}{TD_{50_V}} - 1))}} \tag{1}$$

The depicted data points show specific cardiovascular events in patients after treatment with corresponding dose constraints and cardiovascular toxicity value in the great vessels. Even though tumors may be killed, if unnecessarily high doses are delivered to vital structures, long term health complications can ensue. The model above considers only dose constraints as a factor in cardiovascular toxicity. Therefore, this model of cardiovascular toxicity has limited predictability in terms of complications due to treatment. This is due to the fact that several other patient risk factors can contribute to cardiovascular toxicity

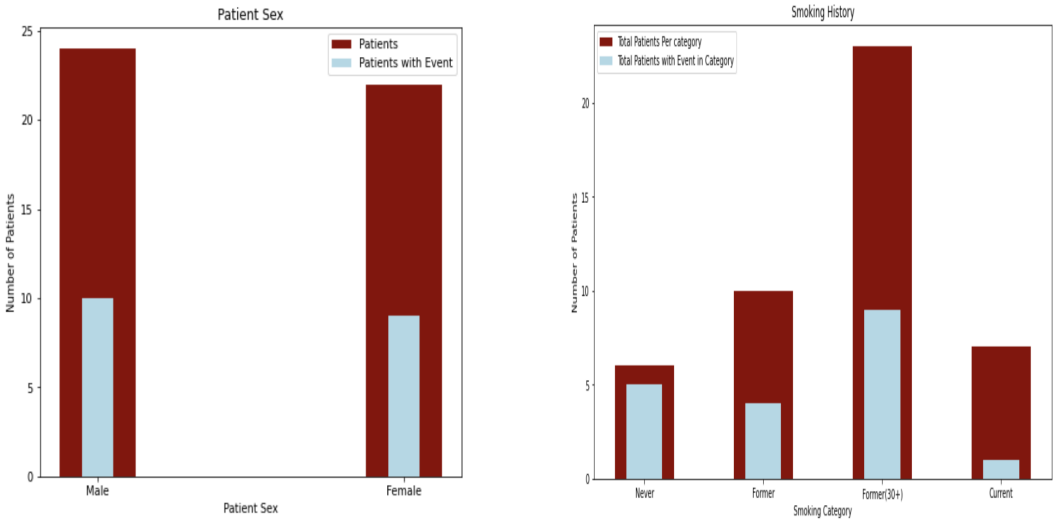
and overall risk and thus potential events. In this paper, we explore the possibility of using machine learning techniques to glean lower level relationships between many different patient and treatment related parameters in addition to just organ-related NTCP.

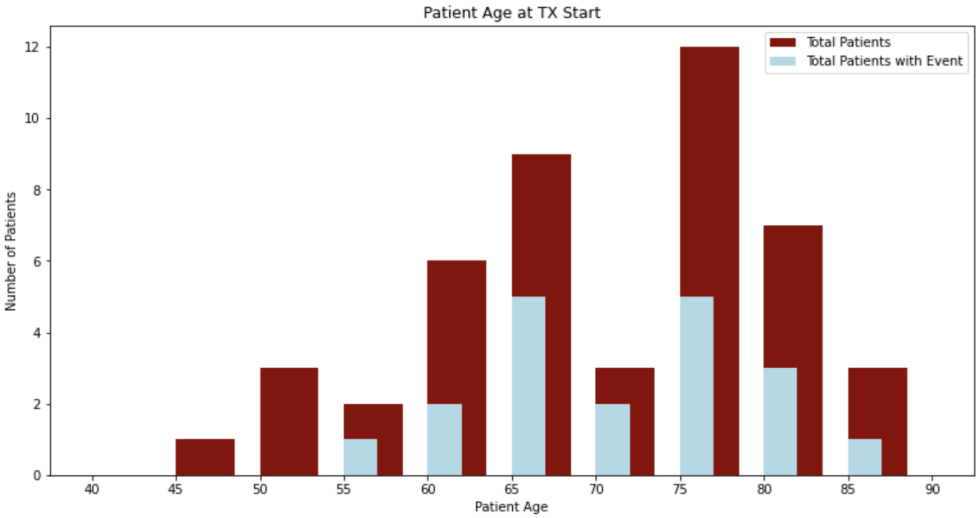
In Equation 1, each organ related NTCP is expressed as a log-logistic function of the specified treatment dose parameter. This NTCP value is calculated for each organ with respect to each of the 5 identified dosimetry statistics: D0.5cc, D1cc, D4cc, Dmax, and V25Gy.

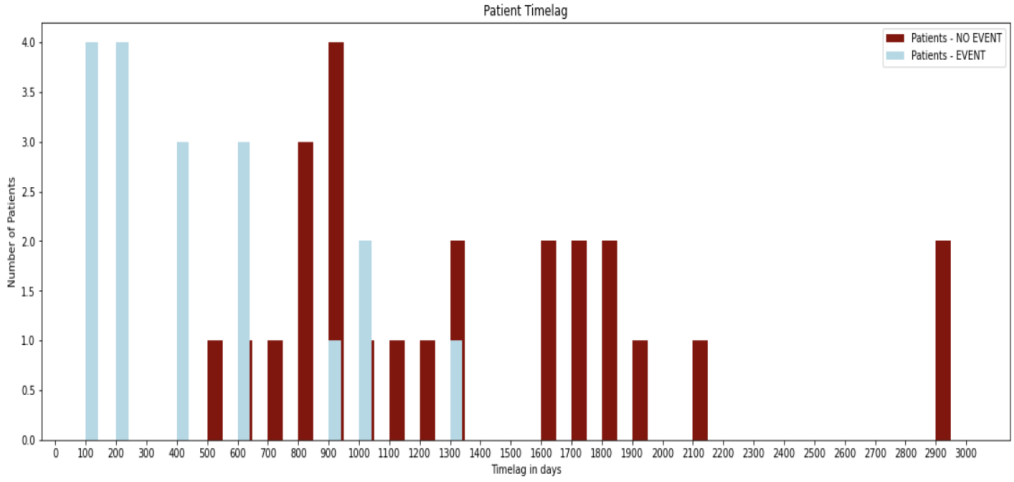
As mentioned above, the model proposed in this paper will consider more patient related parameters in the prediction. These additional parameters are: Patient Age, Patient Race, Patient Smoking history, General Tumor Location, Tumor position(X, Y, Z), Number of prior treatment courses, and Time-lag. We define patient age to be considered at the date of the start of treatment. Smoking history in our model is defined in 4 categories: Never Smokers, Former Smokers, Former smokers who quit after over 30 years of smoking, and Current smokers. General tumor location is defined in our model from the initial medical reports detailing abstract location, while tumor position is determined using the NTCS system after the treatment. Time-lag is defined in our model as the time between the start date of the treatment and the date of an event if one occurred. In the case that an event did not occur, the time-lag of the patient is the time between the treatment start and either the patient's day of death, or the current date.

In our cohort, we had data from 46 different patient's treatments. These patients suffered from several different types of thoracic cancers and had an assortment of complications post-treatment. Out of all 46 patients, 19 patients experienced some form of complication after treatment, while the remaining 27 experienced none. The following images display the

breakdown of the patient data-set with respect to each individual feature.







As seen in these images, the data is relatively diverse with respect to each feature. The data set includes about a 50-50 split in male-female patients. Each category of smoking history is well represented, although the category of former smokers for over 30 years includes the most patients. Patient ages range from about 45 to 90 years old with many distributed heavily in between. The range of patient time-lags extends from as little as 100 days to as long as almost 10 years.

Feature	Correlation With Event
Y	0.237121086
heart_v25cc	0.180979599
Z	0.128298772
Age_at_TX	0.12509406
aorta_v25cc	0.034163098
pa_D4cc	0.022687787
aorta_D4cc	0.015509491
heart_D4cc	0.009922539
pa_D1cc	0.007138522
pa_v25cc	0.00643157
vc_max_dose	0.00311592
aorta_D1cc	0.001383608
aorta_D0.5cc	-0.000306508
aorta_max_dose	-0.001386344
pa_D0.5cc	-0.006108367
Sex	-0.007685716
vc_v25cc	-0.009986798
vc_D0.5cc	-0.01439533
heart_D1cc	-0.026302917
vc_D1cc	-0.029697644
pa_max_dose	-0.033254684
heart_D0.5cc	-0.04067346
X	-0.043570589
heart_max_dose	-0.051855583
vc_D4cc	-0.087971438
Prior_Courses	-0.164771156
Timelag	-0.630383557

This table shows the statistical correlation between whether or not a patient had an event and each of the other features in this data set. As seen in this table, time-lag is most strongly correlated with having an event. Furthermore, other strongly correlated features include the Y and Z coordinates of the tumor, V25 to the heart, number of prior courses, and Age.

2 Methodology

The aim of this project is to apply machine learning to specific SBRT plans and model the overall risk posed to patient. We plan to generate a multi-dimensional model of patient risk from organ-specific cardiovascular toxicity from the SBRT based on: 1) dosimetric characteristics of the treatment plan, 2) elapsed time between RT treatment and event of interest and 3) patient specific risk factors. A machine learning model will therefore be trained to determine risk using the following inputs: Organ-specific V25Gy, D0.5cc, D1cc, D4cc, and Dmax from the SBRT plan; 3D coordinates of the maximum dose point in the organ (determined using the NTCS system); the presence or absence of cardiovascular risk factors such as smoking history, prior cardiac events, etc. ; and patient specific criteria such as age, gender, race, and genetic mutations/variations.

We will utilize two independent machine learning approaches to help develop this model: 1) a RandomForest Classification model (RF) 2) a Support Vector Classification model (SVC), both of which have been proven to successfully classify multidimensional data. Each model will be trained to output a prediction of the class probability whether or not the given SBRT plan will result in a cardiovascular event in the patient. Effectively, the positive class probability will represent the risk to the patient associated with the treatment plan. Thus we define patient $RISK_{OVERALL}$ as the probability of the positive class in the binary classification problem. This can be interpreted as a function of the following characteristics associated with the patient P, and the specific treatment plan T.

$$RISK_{OVERALL} = F(P_{Timelag}, P_{TumorPosition}, P_{Characteristics}, \sum_{j=1}^{4}\sum_{i=1}^{5}(A_{ij} * T^{(j)}_{NTCP_i})) \quad (2)$$

In this interpretation of $RISK_{OVERALL}$, $P_{Timelag}$ is the time-lag associated with patient P, $P_{TumorPosition}$ includes the coordinates of the tumor in patient P and the general category of tumor location, and $P_{Characteristics}$ includes the remaining patient characteristics such as age, sex, smoking history, and prior treatments. Furthermore, j ranging from 1 to 4 indicates which of the 4 specific organs (heart, aorta, vena cava, pulmonary artery) and i ranging from 1 to 5 indicates which of the dosimetry statistics specific to that organ is considered (D0.5cc, D1cc, D4cc, Dmax, V25Gy). Therefore $T^{(j)}_{NTCP_i}$ represents the Normal Tissue Complication Probability estimated by statistic i on organ j in the given treatment plan T and the corresponding A_{ij} is a scaling coefficient.

The machine learning problem posed here boils down to the supervised learning problem of binary classification and more specifically the prediction of binary class probabilities. For any given patient, treatment pair (P, T), we are trying to predict the probability of whether that patient will experience a cardiac event as a result of the treatment. We denoted the binary class 1 to be the class of patient experiencing an event and 0 for no event. Thus, we can model risk of an event as the probability of the given (P, T) falling in class 1.

Due to the relatively small sample size in our cohort, training and testing will be performed using five-fold cross validation. In the training of each independent model, the patients will be divided into training and validation groups (which will be used to perform model fitting and evaluation, respectively) over five distinct iterations in which different groupings are used each time. Repeating this process with each patient subset taking a

turn in the validation set allows for robust evaluation of the model accuracy while avoiding over-fitting as much as possible. For each patient, the toxicity data points calculated by the NTCP equation were input to the model along with other relevant patient information including patient age at treatment, time lag between treatment and cardiovascular event, and patient specific risk factors.

In future efforts, we can improve the accuracy of our model by a larger sample of more detailed data. The risk of a given treatment is an extremely complex problem and there are certainly influencing factors other than the ones listed here. On a data-set including many more patients, we could improve our prediction by collecting genome-wide genotype data on DNA specimens to be collected from the participants.

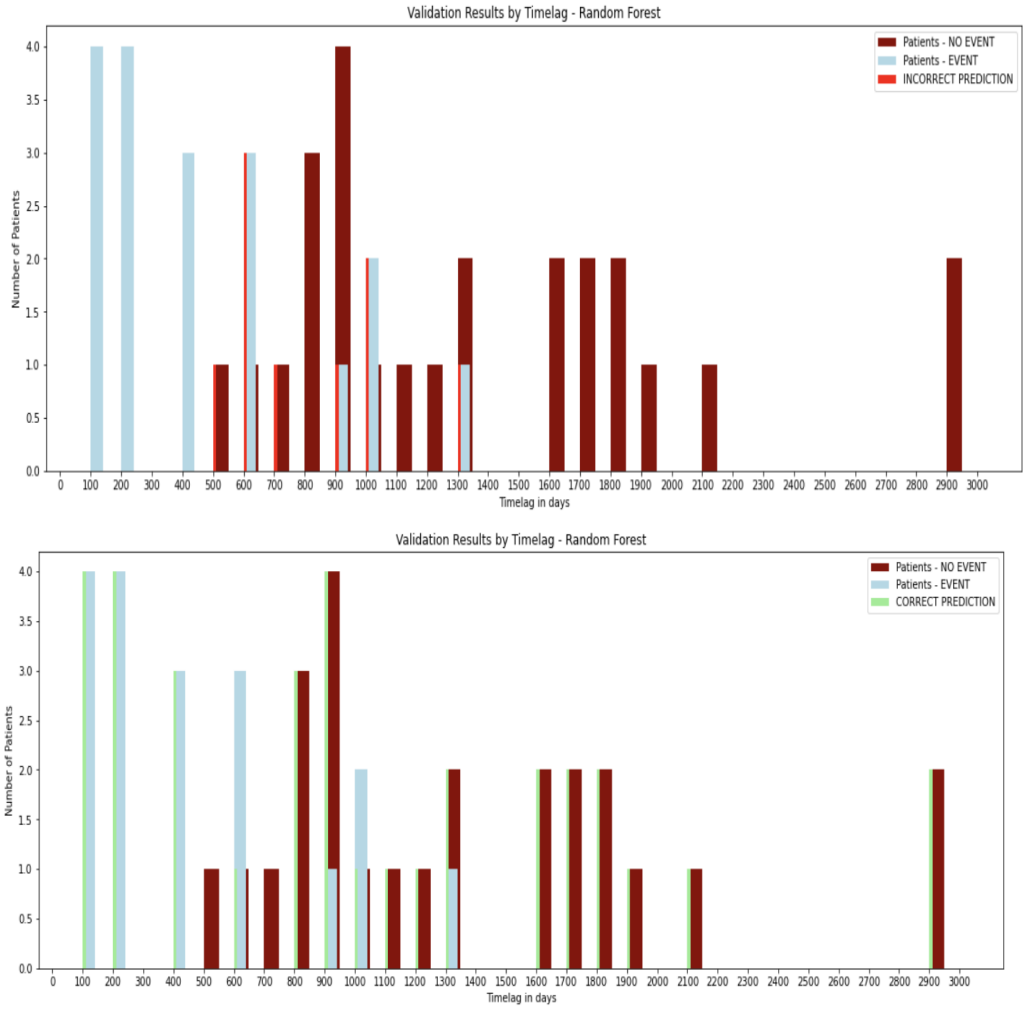
3 Classification

Initially, given the features described above, we started with a model configuration to be trained on all of the input features, except patient time-lag. We performed a grid search cross validation to compute the best set of hyper-parameters for the RF model. Using 5 fold cross validation, we iterated over 50 different combinations of possible hyper-parameters to use in the model. These parameters included the number of estimators used by the RF, the max depth of each individual decision tree, and the max number of features used in each decision tree training. We found that, in the case of this set of patients, the best model hyper-parameters were 100 individual estimators in the ensemble, a max decision tree depth of 2, and at most 2 of the features used in each decision tree.

To evaluate this initial model configuration, we again used the 5 fold cross validation.

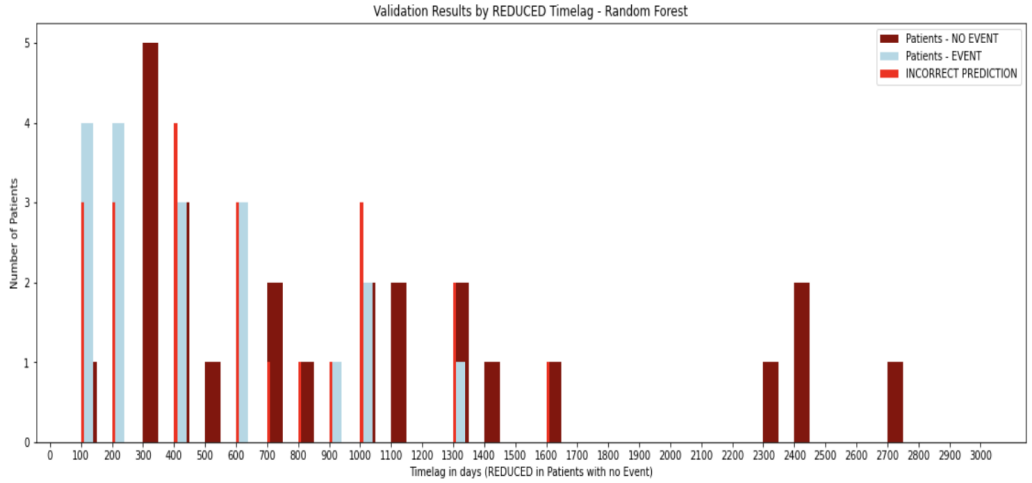
The set of 46 patients was split randomly into 4 groups of 9 and 1 group of 10. During each of the 5 folds, a new RF model, with the optimal parameters found in the grid search, was trained exclusively on 4 out of the 5 groups of patients. Then, we used the model to predict the unseen group. At each fold, we saved these validation predictions for model evaluation. Effectively, after all 5 folds, each patient had been predicted by a model that had not seen that specific patient during training. Thus this provided a more robust way of evaluating the model. At the end, we trained a final model on all of the patients and evaluated the accuracy on those same training patients. We found that, during validation, the model achieved an accuracy of 52.2% on all of the validation predictions and an overall accuracy of 73.9% when trained and evaluated on all patients.

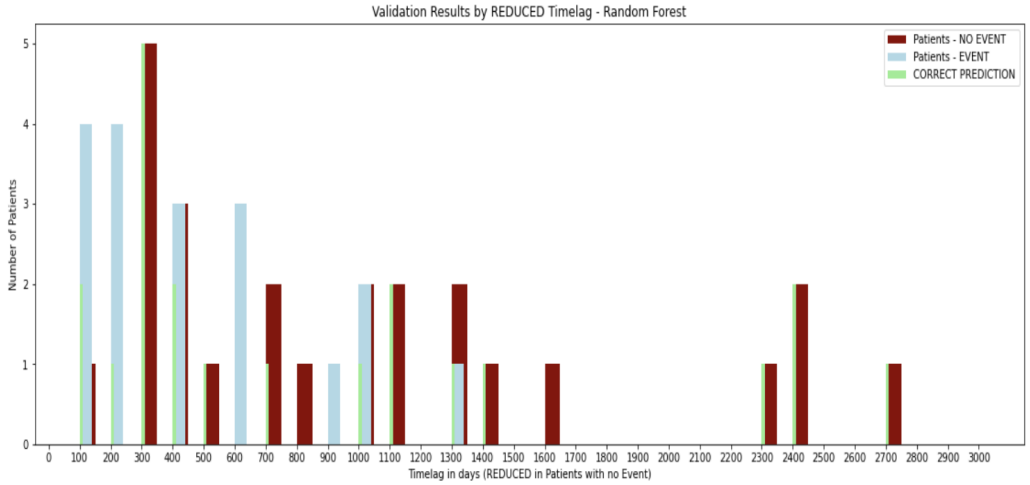
Given that the model did not perform well on the set of patients, we decided to compare the model's performance if time-lag was included as a feature. Again, we performed the grid search with 5 fold cross validation but found the new best parameters. With time-lag, the optimal parameters for the RF model were 50 estimators in the ensemble, a max decision tree depth of 2, and all features used in each decision tree. Using these new optimal model hyper-parameters, we performed the same 5 fold cross validation process to evaluate the performance of this model. This time, we observed a much better prediction accuracy. The model achieved an accuracy of 80.4% on the validation predictions and an accuracy of 91.3% when trained and evaluated on all the patients.



As shown in the image above, it is clear that there is a strong relationship between the absence of a patient cardiac event and a large time-lag. Furthermore, because of this strong correlation, the RF model is relying on a patient's time-lag to make its classification prediction. All of the model's mispredictions at this point occur in the intermediate range of time-lags. As shown in the image, all of the patients with long time-lags and short time-lags are predicted correctly. Recall that we define time-lag in this model as the time between the start of a patients treatment and the date of an event. However, in the case of the absence of an event, either the current date or the patient's date of death was used.

As an attempt to correct some of these mispredictions, we tried reducing the time-lag in each patient who did not experience an event. This effort was made to reduce the disparity in time-lag between patients in each category and thus reduce bias in the model so that relationships between other features could be identified. The same cross validation process was applied, and the results for each validation patient were recorded. This time the same model achieved an accuracy of 56.7% on the validation predictions and an overall accuracy of 89.1% when trained and evaluated on all the patients.





These same three variations of data were also used for prediction with the SVC model

instead of the RF model. With the SVC, we found that using all of the features yielded an overall accuracy of 89.1% on the same patients used for training but an accuracy of only 60.7% on all of the validation predictions. When time-lag was decreased in patients with no event, the accuracy again went down to 80.4% on the training patients and 43.5% on the validation predictions. When time-lag was removed the model predicted the training patients with an accuracy of 84.7% and its validation predictions had an accuracy of 45.6%

Based on these results, the SVC clearly performed worse than the RF model. This is most likely due to the inherent differences between the RF and SVC models. Essentially, the SVC works to maximize the high dimensional margin between each class while the RF works to split the data set by individual features many times. Since the RF is an ensemble of many underlying decision trees, it is able to better classify these patients. On the other hand, the SVC tries to maximize the margin between classes but in the case of these patients, the "margins" overlap too closely for it to yield a good prediction.

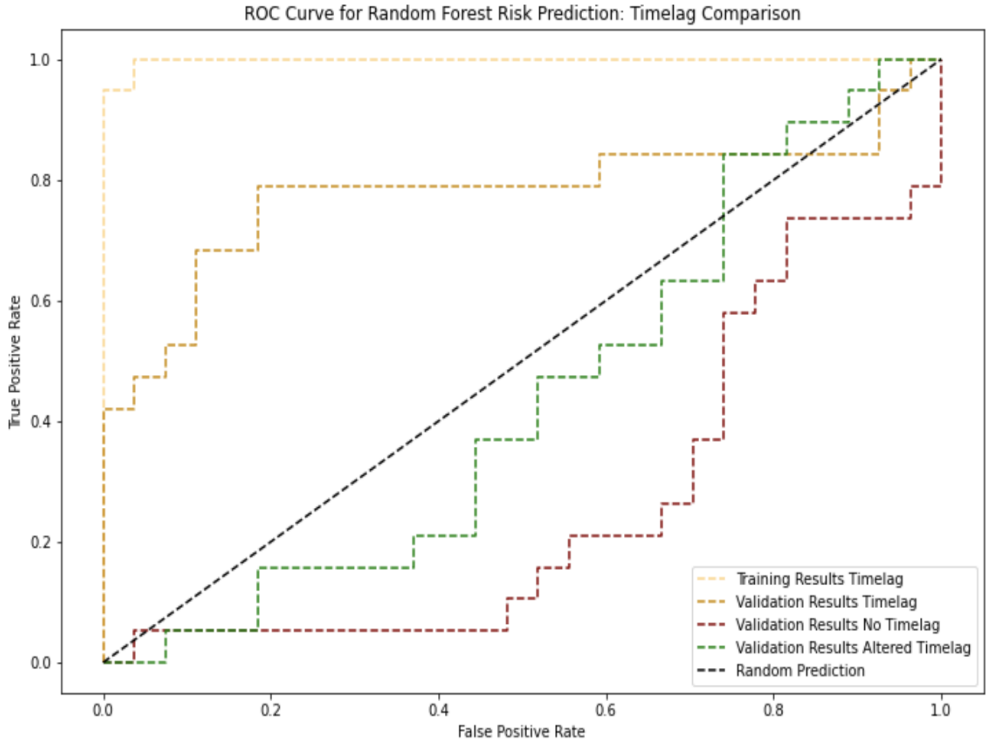
Clearly, the model relies on time lag to make its predictions given this relatively small set of data. However, in order to make useful predictions, the model would need to operate without having access to time-lag in advance. Thus, we decided to switch perspectives and explore the possibility of modeling time-lag itself. That is, given all the same patient info and whether or not the patient experienced an event, how long would the patient have until experiencing an event. Using a Random Forest regression model, we performed a regression to predict time-lag. Using the same 5 fold cross validation process described above, we formulated time-lag predictions for each patient in the validation set during each fold of cross validation. In the end, we evaluated the model using the root mean squared error(RMSE) between the prediction of time-lag, and the time-lag itself. This was done in two separate

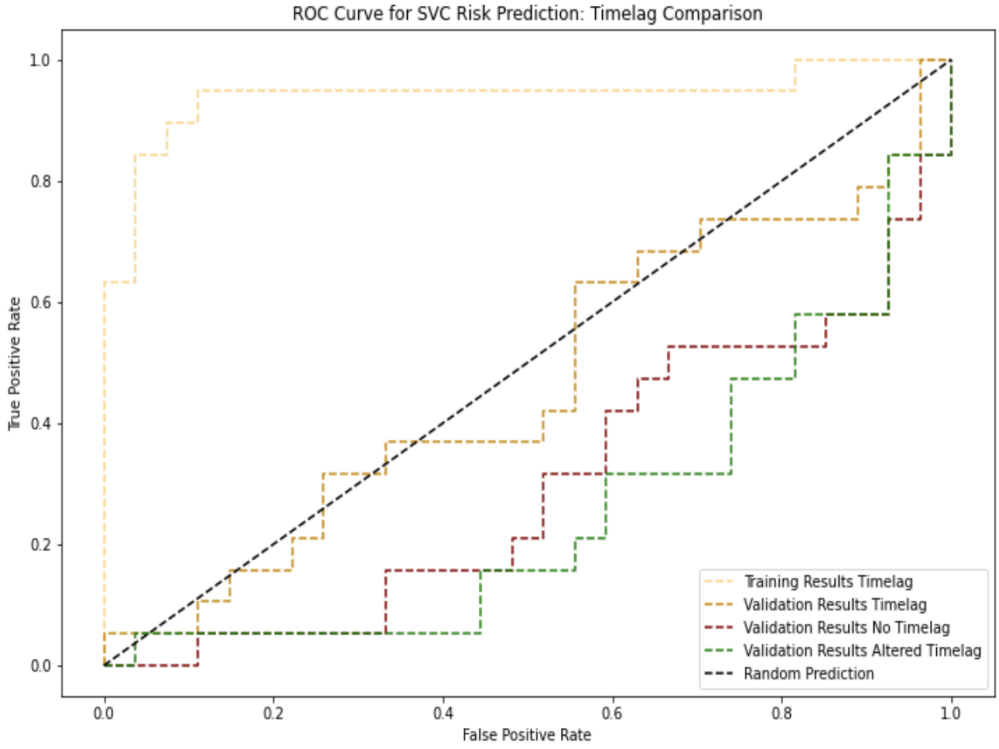
ways. First, separate from the cross validation, the RF regression model was trained on all of the patients and the RMSE was calculated for the predictions on those same training patients. This yielded an RMSE of about 0.7 years. Then, the RMSE was calculated for each time-lag in the validation set. This yielded an RMSE of about 2.1 years.

4 Discussion

After analyzing the results of the risk predictions using both the Random Forest model and the Support Vector model, there are some clear takeaways. First, in the case of both models, there is significant over fitting with respect to the training data. In all cases, the accuracy of the predictions on the training data is much higher than that on the validation data. For the most part, this is to be expected on a data set of this size. When trained on around 35 patients, this complex model is only partially able to determine the underlying relationships between each feature.

Furthermore, it appears that the models are biased due to the features that we are giving it to make predictions. The occurrence of an event is strongly correlated with the patient's time-lag. Given how we compute time-lag, this makes perfect sense. Patients with no event, will have a long time-lag, while patients with an event will have a comparably short time-lag. Because of this correlation, we believe that the model is weighting time-lag much too heavily and thus introducing bias into the prediction. When we tried to combat this by both separately altering and removing time-lag as a feature, the predictability of the model decreased significantly.


ROC Curve for Random Forest Risk Prediction: Timelag Comparison
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Training Results Timelag
Validation Results Timelag
Validation Results No Timelag
Validation Results Altered Timelag
Random Prediction

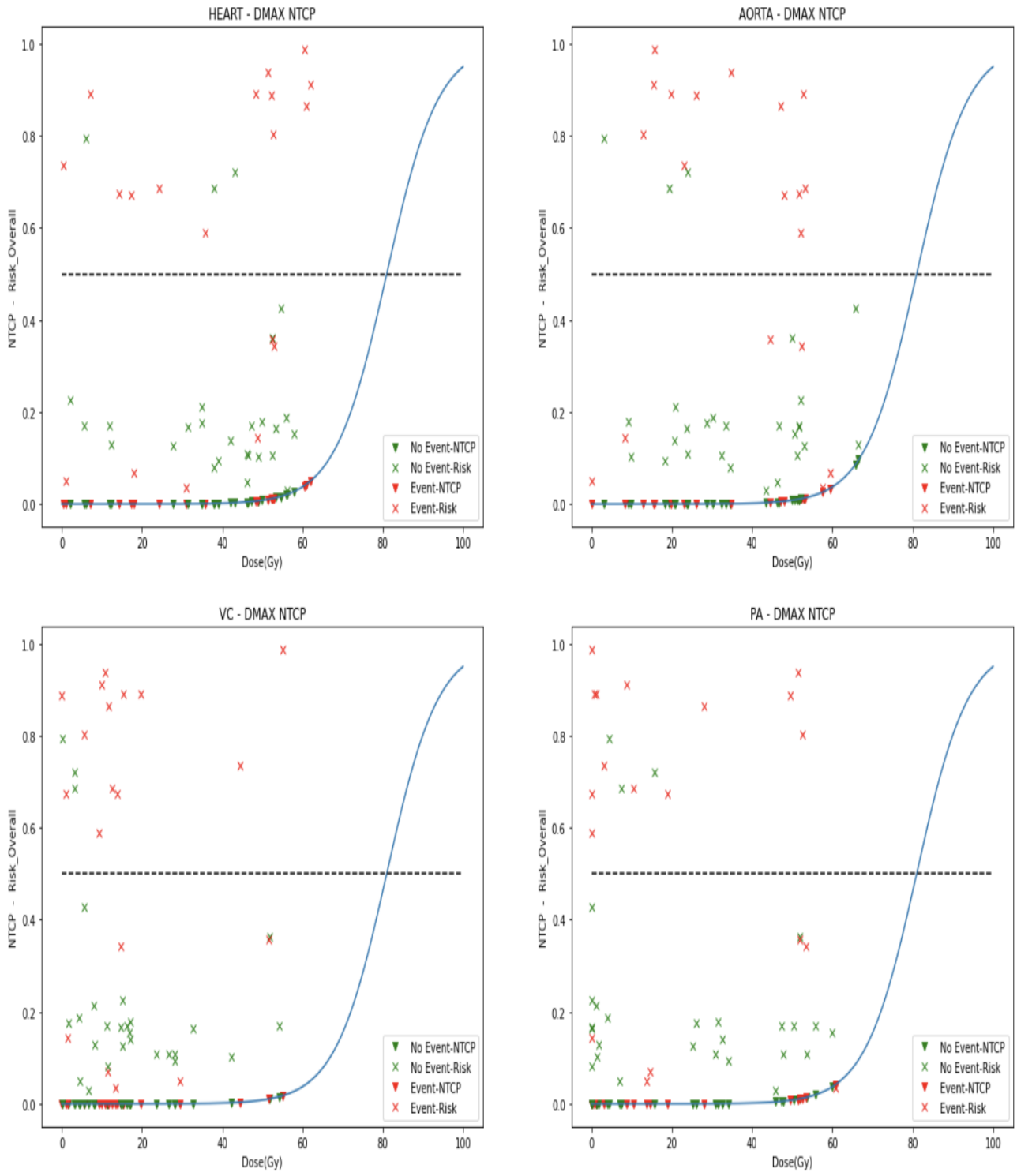

ROC Curve for SVC Risk Prediction: Timelag Comparison
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Training Results Timelag
Validation Results Timelag
Validation Results No Timelag
Validation Results Altered Timelag
Random Prediction

This plots above demonstrate the ROC curve for the classification in 4 cases for both the RF model and the SVC model. The orange colors show the predictions after including all features including time-lag. The light orange represents the predictions on the same patients used for training while the dark orange shows the validation predictions. Finally, the dark red and green show the predictions on the validation data when trying to combat the bias associated with time-lag. Green represents the predictions after reducing the time-lag in all patients with no event and red represents the predictions after removing time-lag completely.

Furthermore, we considered the results of the regression on time-lag. Given that time-lag is by far the most important feature in terms of the predictability of the model, and that time-lag is not a feature we can ascertain before treatment, the regression may be helpful. The results of the regression showed that we can predict time-lag with an RMSE of about 2.1 years on the unseen validation data. This is not incredibly accurate given that the average patient time-lag is 3 years. However, we may be able to use the output of this regression as the predicted time-lag and thus as an input to the risk prediction model. This could yield some interesting results moving forward. However, it still poses a problem given that this time-lag prediction is determined using whether a patient had an event as a binary feature.

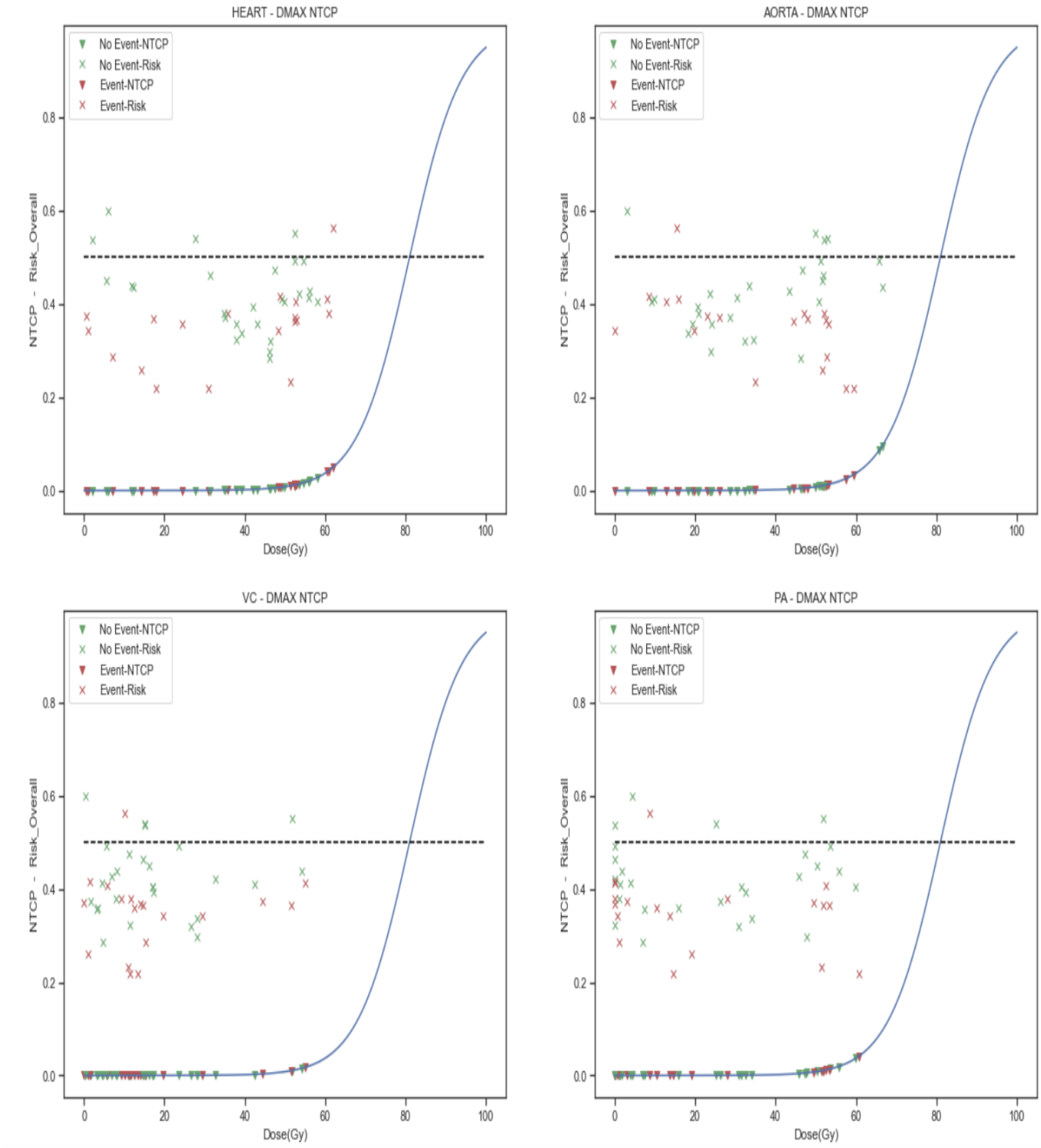
Aside from the time-lag concerns, this machine learning approach yielded some interesting results in terms of the predictability of patient risk. As discussed in the introduction, we had previously considered the possibility of modeling using exclusively NTCP. In some cases, NTCP could accurately identify risky patients, but in most patients this approach was inconclusive.

These images detail a comparison between the predictions of the RF machine learning model and the NTCP associated with each organ and the max dose to that organ during



treatment. Each X represents a prediction from the RF model while the triangles show the corresponding NTCP of the patient. Green indicates no event while red indicates an event.

As shown in these images, in relatively low doses, the machine learning approach is able to do well in the classification. However in other cases the NTCP model better represents the risk.



As a comparison, these images show the RF model predictions after being trained on the data without patient time-lag. Its clear from these images that the model is mispredicting almost all of the patients who experienced an event.

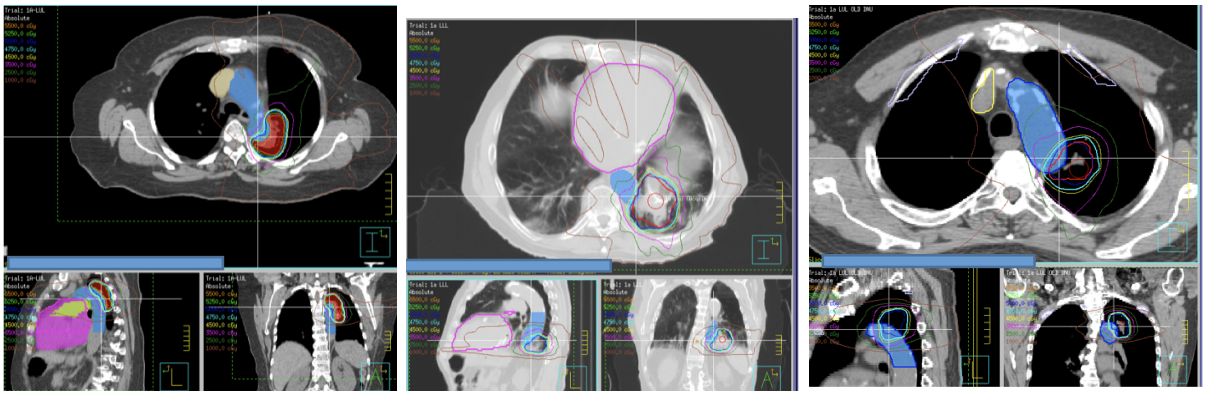
	Event	prediction	heart_v25cc	heart_max_dose	aorta_v25cc	aorta_max_dose	pa_v25cc	pa_max_dose	vc_v25cc	vc_max_dose
HM1	0	0.540434	0.265625	27.666470	40.187500	52.931055	0.015625	25.142346	0.000000	15.130018
WJ	0	0.537814	0.000000	2.082134	8.609375	52.070930	0.000000	0.000000	0.000000	15.146054
SB	0	0.551146	36.046875	52.418956	8.609375	49.953960	1.203125	51.780766	5.546875	51.812725
PG	0	0.599034	0.000000	6.056546	0.000000	3.131939	0.000000	4.382165	0.000000	0.270842
DJ	1	0.379739	19.421875	35.666619	31.921875	52.115100	0.000000	0.000000	0.000000	9.300355
IB	1	0.416666	4.687500	48.680009	0.000000	8.373321	0.000000	0.000000	0.000000	1.412482
FJ1	1	0.342151	4.750000	48.365220	0.000000	19.816527	0.000000	0.596838	0.000000	19.675293
HD	1	0.370099	9.421875	52.337116	0.031250	25.979862	7.250000	49.484763	0.000000	0.000000
AJ	1	0.364028	47.265625	52.500380	10.718750	44.559248	2.656250	51.934535	9.453125	51.651730
BL	1	0.378523	13.625000	60.918565	2.968750	47.081517	1.390625	28.063299	0.000000	11.697476
GB	1	0.233035	12.312500	51.391299	1.328125	34.841891	4.859375	51.391299	0.000000	10.905179
WL	1	0.287305	0.000000	7.069553	9.718750	52.784173	0.000000	1.195615	0.000000	15.366442
MJ1	1	0.217764	0.000000	18.014192	8.281250	59.407711	0.000000	14.569057	0.000000	11.454040
RD	1	0.364797	44.593750	52.851391	42.484375	52.413933	4.328125	53.415919	0.000000	14.701987
LL	1	0.343363	0.000000	1.090614	0.000000	0.094535	0.000000	13.725299	0.890625	29.529727
DE	1	0.259519	0.000000	14.347536	35.468750	51.694292	0.000000	19.008777	0.000000	1.105028
MK	1	0.406679	23.703125	52.642059	0.000000	12.860631	12.375000	52.551757	0.000000	5.733456
MM1	1	0.358363	0.000000	24.366144	12.375000	53.123147	0.000000	10.443853	0.000000	12.488627
GV	1	0.412287	19.859375	60.471948	0.000000	15.761207	0.000000	0.000000	2.078125	55.076543
BC	1	0.374527	0.000000	0.513207	0.000000	23.076514	0.000000	3.023778	8.750000	44.463597
DJ1	1	0.369334	0.000000	17.233691	13.125000	47.951575	0.000000	0.000000	0.000000	13.932511
HI	1	0.218333	1.109375	31.011604	39.609375	57.496405	13.890625	60.819327	0.000000	13.377234

The table above demonstrates the exact patients on which the model produced an

incorrect prediction. The prediction column shows the validation predictions of the patient's *Risk_Overall* from the RF model trained without time-lag. Upon examination of the key dosimetry statistics associated with the treatment of these patients, it appears that the patients who experienced an event had a significantly high dose to at least one of the given organs.

5 Future Improvements

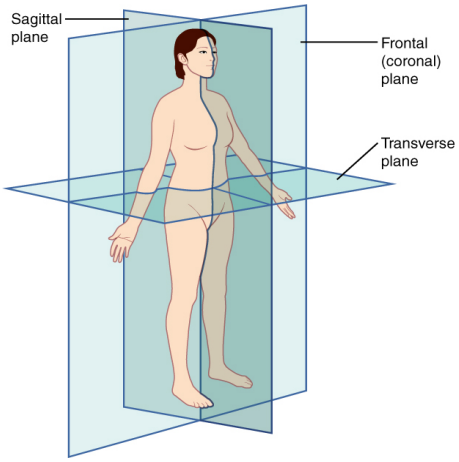
Data suggests that there could be a location dependence to toxicity for a given organ. For example, high doses to the aortic arch do not lead to cardiovascular events even with 53.0Gy (HM1) as shown here in the rightmost image.



Patients: DE (Event), DJ (Event), and HM1 (No Event)

On the other hand, high doses to the descending or abdominal aorta can lead to cardiovascular events even with 51.5Gy (DJ, DE) as shown in the left and middle images respectively. These patients pictured above all received doses in excess of 50Gy to the aorta. However, only two of them experienced a cardiac event as a result. Moving forward, we must take into consideration more specific dose locations using a normalized coordinate system.

Additionally, given more time and data, this model could be extended into problem of multi-class classification. The binary classification model discussed in this paper over-simplifies the true nature of medical complications. Instead of a generic Event/No Event prediction, we could model specific types of events. This multi-class modeling process would be much more complex, and therefore require much more data in order to make accurate predictions.



Furthermore, as discussed above we observed relatively high correlation between a patient having an event, and both the Y and Z coordinates of the tumor. From the image above, the Y axis corresponds to the saggital axis, and the Z axis corresponds to the axis along the spine. At this point, we are using non-normalized values of these coordinates. However, as discussed in the Wang et al. paper(3), due to deviations in patient thoracic size, build, and general organ location, these coordinates could represent very different areas in different patients. Using a normalized coordinate system similar to the NTCS(Normalized Thoracic Coordinate System) proposed in the Wang et al. paper, these coordinates could prove to be much more useful in the predictive model.

These attempts have shown very interesting results and the machine learning approach

shows lots of promise. Moving forward, the best way to improve results and reduce the bias introduced by time-lag is to get much more data. This is an extremely complex problem, and one that a machine learning algorithm is best suited to solve. However, it is very hard to create a model of this complexity with fewer than 50 patients to work with.

References:

(1) Xue J, Kubicek G, Patel A, Goldsmith B, Asbell SO, LaCouture TA. Validity of Current Stereotactic Body Radiation Therapy Dose Constraints for Aorta and Major Vessels. Seminars in Radiation Oncology. 2016.

(2) Bentzen SM, Tucker SL. Quantifying the position and steepness of radiation dose-response curves. International Journal of Radiation Biology. 1997.

(3) Wang H, Bai J, Zhang Y. Normalized thoracic coordinate system for atlas mapping in 3D CT images. Prog Nat Sci. 2008;

(4) Darby SC, Ewertz M, McGale P, Bennet AM, Blom-Goldman U, Brnønum D, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med. 2013;

(5) Kataria T, Bisht SS, Gupta D, Abhishek A, Basu T, Narang K, et al. Quantification of coronary artery motion and internal risk volume from ECG gated radiotherapy planning scans. Radiother Oncol. 2016;

(6) Wennstig AK, Garmo H, Hållström P, Nyström PW, Edlund P, Blomqvist C, et al. Inter-observer variation in delineating the coronary arteries as organs at risk. Radiother Oncol. 2017; (7) Khera A V., Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics. 2018.

(8) Jonathan Colen