Predicting Erectile Function in Patients Undergoing Stereotactic Body Radiation Therapy

09688532 and 91455724

2017-01-26

Abstract

Stereotactic Body Radiation Therapy (SBRT) is a new method for treating prostate cancer. This study looks at the potential for SBRT patients to experience erectile dysfunction, an important side-effect that is associated with the treatment. Using a logistic regression model, this study finds that baseline variables including age, score on a quality of life survey, erectile function at baseline, and prostate-specific antigen levels can be used to predict the probability of having erectile function two years after receiving SBRT. The model performs well, with a value of 0.73 area under curve (AUC). Following the results from the model, we conclude that around 38 percent of patients who undergo SBRT are likely to have erectile function at two years, compared to 49 percent at baseline.

Introduction

Stereotactic Body Radiation Therapy (SBRT) is a new treatment option available for prostate cancer, the most common form of male cancer. The treatment is characterized by relatively few instances of high dose radiation over a short period of time, as opposed to previous treatment methods with small doses of radiation over a longer period of time. Given that SBRT has around the same tumor kill effectiveness as the previous treatment, and that prostate cancer is quite common and very treatable, the largest concern associated with treatment selection is the lifestyle impact of side-effects. In particular, the most relevant side-effect associated with SBRT is the potential for erectile dysfunction.

In this study, data collected from SBRT patients is used to fit a logistic regression model exploring the likelihood of losing erectile function following treatment with SBRT. The data included baseline variables such as: age; Gleason score, the grade of the patient's cancer; prostate-specific antigen (PSA), a continuous measurement from a blood test associated with aggressiveness of one's cancer; T-Stage Group, categorizing

the size and spread of the tumor; androgen deprivation therapy (ADT), a variable indicating whether or not the patient was chemically castrated; body-mass index (BMI), a measure of height-to-weight ratio; and erectile function. A more detailed look at our variables can be found in Table 1, which includes descriptive statistics for each variable, including the mean, standard deviation, median, minimum, and maximum.

Our model shows that age, score on a health survey, and prostate-specific antigen levels can be used to predict the probability of erectile function two years post-treatment with reasonably high accuracy.

Methods

Clinical and biomarker data were collected from 325 patients with prostate cancer that experienced SBRT and were followed for 2 years or longer between 2008 and 2013. All the patients in the data were treated at a single institution. The dataset does not have any missing values, but only contains information from patients who underwent radiation therapy. These radiation therapy patients are generally older, sicker, and at more advanced stages of cancer than patients who elect for surgery. The dataset also contains information from patients who already reported no erectile function at baseline, and these patients were not removed before analysis. Table 2 contains counts of patients by erectile function at baseline and at two years.

To check for outliers or distribution oddities, we plotted a histogram for each of the continuous variables. If we found outliers in a histogram, we would first look at those outliers to determine if their values are theoretically possible. If the outliers were clearly typos, we would contact the data provider and then remove them. Any values that were strange but theoretically possible were left in the model. In particular, subject 67 has a BMI of 112, which is very different from the other BMI values but still within the realm of possibility.

Collinearity was another potential issue, which would not reduce the predictive power of the model overall, but could reduce the preciseness of the calculations of individual predictors. Hence we calculated the correlation between each of the potential predictors and visualized the correlation matrix using the corrplot function in R, so that we could directly find any pair of predictors that have high correlations. In our dataset, we expected HRQOL and Erectile function at Baseline to be highly correlated, because the HRQOL survey includes questions that ask about erectile function. We found that the correlation between these two variables was 0.73, but kept both variables in the model because they were both shown to be highly significant, as explained in more detail below.

Logistic Regression

We built a logistic regression model to predict the probability that one will have erectile function at two years after SBRT. The model was built in the R language version 3.3.2 (2016-10-31), using the package glm. See appendix XX for the full reproducible code document.

Our model was fit with "Erectile Function at 2 Years" as the outcome variable, and all other variables (Age, Gleason Score, PSA, T-Stage Group, ADT, BMI, Erectile Function at Baseline) as predictors. Age in the model was centered by subtracting the mean age from each entry. Following a recommendation from the physician, we collapsed Gleason score into "low" (GS <= 6), "medium" (GS == 7), and "high" (GS >= 8) categories. Because there was only a single patient with T-Stage group >= 2, we collapsed T-Stage group into two categories, 0 and 1, adding the patient with T-Stage group >= 2 into the 1 group.

We decided to use logistic regression because it meshed with the objectives of the study: classification, interpretation, and prediction. Logistic regression is well-suited for our outcome variable "Erectile Function at 2 Years," which is a binary variable, allows us to estimate the probability of having function at 2 years, and also enables the straightforward interpretation of the effects of each predictor in the model. We briefly considered using other classification methods, such as a random forest classifier or a support vector machine classifier, but since it would be hard or impossible to interpret the effects of each predictor using those models, we decided to focus on logistic regression. Besides our final model, we also tried a few other logistic regression models including interaction terms, in particular between age and other covariates. These models did not perform noticeably better than our reported final model. Another model, fit using only data from patients who reported erectile function at baseline, also did not show any improvement.

To analyze the possible multicollinear effects of including both HRQOL and Erectile function at baseline variables in the model, we fit the model with each variable individually, and then with both. We decided to keep both variables in the model because both variables were shown to be highly significant when fit together, and because models without both variables performed worse.

We evaluated the performance of our model based on its prediction accuracy rate, goodness of fit, and the area under its receiver operating characteristic (ROC) curve. We calculated the prediction accuracy rate as the proportion of subjects that we predicted correctly among all 325 subjects when using a probability of 0.5 as the threshold. We tested the goodness of fit by using the Hosmer-Lemeshow test, which is a commonly used statistical test for goodness of fit for logistic regression models. The Hosmer-Lemeshow test resulted in a X-squared value of 6.8019 and a p-value of 0.5581, meaning we fail to reject the null hypothesis of poor model fit and fail to find any evidence of lack of goodness of fit. We created the ROC curve by plotting the true

positive rate (TPR) against the false positive rate (FPR) at various threshold settings, which could compare the performance of different binary classifiers, and found the area under curve (AUC) to be 0.73. When classifying patients with their final outcome, we used a threshold of 0.5, so that if the predicted probability was greater or equal to 0.5, then we predicted that the subject would have erectile function at 2 years.

Results

Table 1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Patient	325	163.00	93.96	1	325
Age	325	69.14	7.15	44	94
Gleason Score	325	6.58	0.62	4	9
T-Stage Group	325	0.16	0.37	0	2
PSA	325	7.72	4.86	0.80	42.80
HRQOL	325	56.08	30.79	0	100
ADT	325	0.10	0.29	0	1
BMI	325	28.88	6.55	19	112
Erectile Function at Baseline	325	0.49	0.50	0	1
Erectile Function at 2 years	325	0.31	0.46	0	1

Table 2: Counts by erectile function

	Baseline	Two years	Count
1	0	0	153
2	1	0	70
3	0	1	14
4	1	1	88

The coefficient results from our logistic regression model can be found in Table 3. The model shows that Age, HRQOL, and Erectile function at Baseline are highly predictive of Erectile function at two years. It also shows that PSA is marginally significant.

The variables PSA, HRQOL, and Erectile function at baseline were shown to be protective of Erectile function at two years, while the variables Age, Gleason score, T-Stage group, ADT, and BMI were negatively associated.

In general, these results are in line with our original hypothesis, excluding the resulting for PSA. Harmful variables such as Age, Gleason score, T-Stage group, ADT, and BMI would be expected to have a negative association with Erectile function at two years. The fact that most variables in the model were not found to be significant is also not very surprising: all variables were collected at baseline, and since all patients

underwent the same treatment, cancer-related variables were not expected to be very predictive of erectile function at two years. The only real disappointment was Androgen Deprivation Therapy (ADT), a variable indicating the presence of chemical castration, which biologically seems to be a potential cause of erectile dysfunction. However, we do not have information on when the patients went through ADT, relative to when they reported their other variables, so we cannot say for sure why we failed to find an association between ADT and erectile dysfunction.

The most unexpected result is that the odds ratio for PSA is greater than 1, indicating that a higher PSA value is actually protective of erectile function at two years, despite indicating a more aggressive and risky cancer. We would need more domain knowledge and data to attempt to explain why this is the case. However, it should be noted that PSA is only slightly significant in our model.

Table 3: Logistic regression results

	For Erectile function at baseline:
	Odds ratios for one unit change (95% CI)
Age	0.943** (0.899, 0.988)
PSA	1.063* (0.994, 1.137)
'Gleason Score'	0.904 (0.523, 1.561)
'T-Stage Group'	0.603 (0.212, 1.715)
HRQOL	1.023*** (1.007, 1.040)
ADT	0.483 (0.148, 1.581)
BMI	0.979 (0.923, 1.038)
'Erectile Function at Baseline'	4.299*** (1.824, 10.132)
Observations	325
Note:	*p<0.1; **p<0.05; ***p<0.01

The ROC curve produced from our model can be found in Figure 1. The area under the ROC curve is 0.842, indicating reasonable model performance. Using the ROC curve, we used the minimum Euclidean distance between the ROC curve and the upper left corner of the graph window to calculate an optimal classification threshold of 0.426. Using this threshold, 38 percent of patients in the study were predicted to have erectile function after two years down from 49 percent at baseline, which represents a 0.225 error rate from the actual results in the data. For comparison, using 0.5 as the cutoff, 30 percent of patients were expected to retain erectile function, representing a 0.206 error rate. In actuality, 31 percent of patients had function after two years. The benefit of using the calculated threshold of 0.426 over 0.5 is it is more optimistic–by classifying more people as retaining erectile function, we increase the true negative rate, and are less likely to scare off patients from opting for the treatment. We investigated the calculated error rate by using 10-fold cross validation, and found the estimated error rate to be around 0.24 at both 0.426 and 0.5 thresholds.

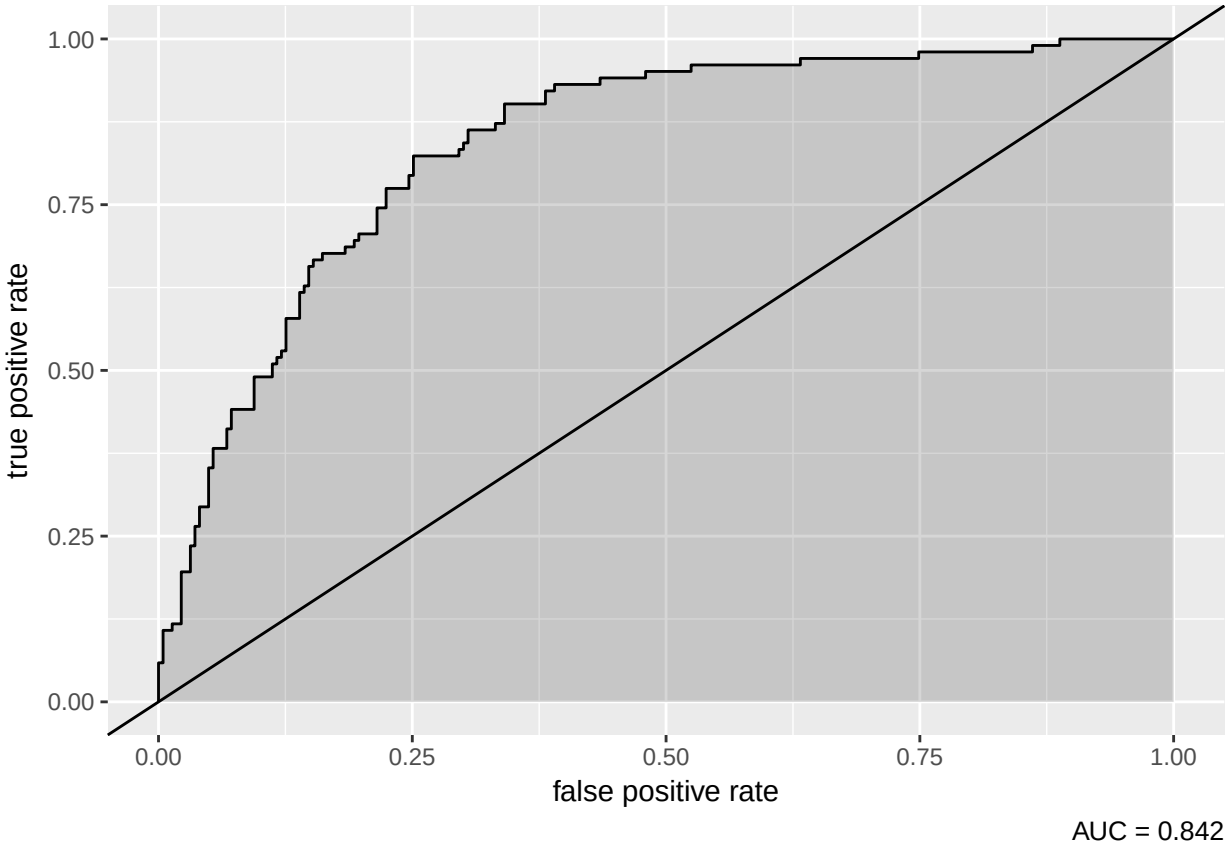

Figure 1: ROC Curve

Conclusions

Based on the logistic regression model, we found that age, HRQOL, and erectile function at baseline were highly predictive of erectile function at two years. The model failed to show that Gleason score, T-Stage group, ADT, or BMI were predictive of erectile function at two years. PSA was found to be moderately predictive, but contrary to expectations, was protective of erectile function. HRQOL and erectile function at baseline were also protective of erectile function, while age was detrimental.

The model was able to produce an ROC curve with good AUC and reasonable predictive probability. Using a cutoff of 0.426, we predicted 38 percent of people to have erectile function at two years, representing a 0.225 error rate when compared to the actual results.

Overall, we conclude that men who are relatively young, have high reported HRQOL scores and PSA levels, and have erectile function at baseline are very likely to retain their erectile function two years afterwards. Of these predictors, HRQOL and age are the most relevant, with PSA levels having less of an impact. Meanwhile, men who are relatively old, have low HRQOL scores, and have low levels of PSA are very likely to lose function. Finally, men who do not have erectile function at baseline are unlikely to gain function back at two years, no matter what their profile looks like.