Survival Prediction Models for Estimating the Benefit of Post-Operative Radiation Therapy for Gallbladder Cancer and Lung Cancer

Jayashree Kalpathy-Cramer PhD[1], William Hersh, MD[1], Jong Song Kim, PhD[2], Charles R Thomas, MD[3], Samuel J Wang, MD, PhD[3,1]
[1]Dept of Medical Informatics and Clinical Epidemiology, and [3]Dept of Radiation Medicine, Oregon Health & Science University, Portland, OR;
[2]Dept of Statistics, Portland State University, Portland, OR

Abstract

The role of post-operative radiotherapy (PORT) is controversial for some cancer sites. In the absence of large randomized controlled trials, survival prediction models can help estimate the predicted benefit of PORT for specific settings. The purpose of this study was to compare the performance of two types of prediction models for estimating the benefit of PORT for two cancer sites. Using data from the Surveillance, Epidemiology, and End Results database, we constructed prediction models for gallbladder (GB) cancer and non-small cell lung cancer (NSCLC) using Cox proportional hazards (CPH) and Random Survival Forests (RSF). We compared validation measures for discrimination and found that both the CPH and RSF models had comparable C-indices. For GB cancer, PORT was associated with improved survival for node positive patients, and for NSCLC, PORT was associated with a survival benefit for patients with N2 disease.

Introduction

Post-operative radiotherapy (PORT) is often recommended after complete surgical resection for certain cancer sites. In some settings, however, the survival benefit of PORT is unclear, particularly for tumor sites where large randomized trials may not have been conducted, or in settings where the results of such studies are equivocal[1, 2].

Gallbladder (GB) cancer is relatively rare cancer with a poor prognosis. Small gallbladder studies have reported a potential benefit of PORT[3], however, there have been no large randomized clinical trials to date, because of the rarity of this tumor.

For non-small cell lung cancer (NSCLC), although several studies have been performed evaluating the use of PORT, results have been equivocal, with some studies suggesting a potential benefit only for certain subsets of patients, such as those with N2 disease, and a detrimental effect for other subsets[4].

Recently published studies[1,2] have demonstrated that regression modeling using Cox proportional hazards (CPH) modeling techniques can help predict which cancer patients may benefit from PORT for GB and NSCLC. While CPH is a commonly used technique for modeling survival data, it requires an assumption of proportionality of hazard rates, and there is suggestion that some of these datasets may not satisfy this assumption. The purpose of this study was to compare the performance of the CPH model with another type of survival prediction model, Random Survival Forests (RSF), a technique that does not require an assumption of proportionality of hazard rates.

Methods

Study Population

The Surveillance, Epidemiology, and End Results 17 (SEER) database from the National Cancer Institute is a large cancer registry covering approximately 26% of the US population[5]. Using the SEER database, we selected two patient populations with resected cancers: GB cancer patients, and NSCLC patients. Patients were selected that were diagnosed between 1988 and 2004. Only patients with a complete surgical resection with microscopically confirmed disease were included. Candidate covariates considered for inclusion were based on known clinically prognostic factors and availability in the SEER database

Statistical Methods

The semiparametric Cox model is one of the most widely used multivariate regression models for censored data. Kaplan-Meier and Cox proportional hazards analyses were performed using the "Design" and "Survival" packages in R, an open source statistical package (http://cran.r-project.org/) .

Although the Cox model is the most widely used for survival analysis, it assumes proportional hazards or that the ratio of the hazards is constant over time. This is often found not to be true in clinical situations. In addition, the Cox model does not reduce to an ordinary linear regression in the absence of censoring. There are many alternatives to the Cox model that do not assume proportional hazards including neural networks, nonlinear accelerated

failure time models, spline-based extensions, ensemble methods like Random Survival Forests and Cox-Aalen models[7-11].

The Random Forests algorithm is a popular machine learning tree method first proposed by Leo Breiman[8]. Randomly drawn bootstrap samples of the data are used to grow the tree. The tree is -then split on randomly selected predictors. The Random Survival Forests were proposed as an extension to Random Forests as a technique to deal with censored data or data for which the exact time of death is not known either because the study participant's status is unknown or the participant is still alive.

We constructed two types of survival prediction models for both the GB and NSCLC cancer sites: Cox proportional hazards (CPH) model and Random Survival Forests (RSF). This analysis was performed using the 'RandomSurvivalForest' package in R[9]. We evaluated the validity of the proportional hazards assumption for both our datasets.

Covariates used in this analysis were age, sex, race, histology, TNM stage and the use of PORT[1]. The continuous variable age was modeled using restricted cubic splines to allow more flexibility in modeling potential non-linear effects. Restricted cubic splines are piecewise cubic polynomials that can be made to be smooth at the 'knots' by matching derivatives. For the CPH model, interaction terms between PORT and stage were also included to reflect the possible effects of stage on the benefit of PORT. We included all variables in the model and did not perform variable selection since this has been shown to improve predictive accuracy[6].

Model performance was internally validated using bootstrapping by measuring discrimination using the concordance index (C-index). This C-index is the probability that given two randomly selected patients, the model correctly predicts which patient will have a better outcome. We also report examples of predicted survival for typical cancer patients using these models.

Results

Gallbladder cancer

A total of 3985 patients met the inclusion criteria for GB cancer. There were 2981 female and 1004 males in the final GB dataset. Of these, 3225 had received PORT while 760 had not received PORT. 931 patients were stage T1, 723 were stage T2, 1255 were stage T3, 269 were stage T4 and 807 were stage M1. 886 were node positive while 1827 were node negative**.**

The CPH model was built first. The C-index for this model is 0.71 using age, sex, race, papillary histology, TNM stage and the use of PORT as covariates. The hazard ratios for the significant covariates are given in table 1. There are significant ($p < 0.05$) interactions between effect of RT and the stage. The hazard ratios of these interaction terms are less than 1 indicating that PORT is associated with improved survival for these population subsets.

	Hazard Ratio	Lower CI	Upper CI
Age	1.02	1.01	1.03
Race=Other	0.86	0.74	1.00
Papillary=Yes	0.56	0.46	0.69
Tstage=.T2	1.31	1.11	1.55
Tstage=.T3	2.32	2.00	2.68
Tstage=.T4	4.60	3.69	5.74
Tstage=M1	5.27	4.44	6.26
zN=N1	1.51	1.35	1.70
XRT=1	1.54	1.14	2.08
Tstage=.T2 * XRT=1	0.64	0.44	0.93
Tstage=.T3 * XRT=1	0.57	0.41	0.81
Tstage=.T4 * XRT=1	0.38	0.24	0.60
Tstage=M1 * XRT=1	0.51	0.33	0.78
zN=N1 * XRT=1	0.73	0.58	0.92

Table 1. Hazard ratios for gallbladder dataset

However, when the proportional hazards assumption was evaluated, we found that there were significant variations in proportionality over time ($p < 0.05$). The covariate with the largest coefficient and most

significance was the use of PORT. This can be seen Figure 1 where there is an initial improvement in survival with PORT as indicated by the separation of the curves in Figure 1 and the negative coefficient in Figure 2, but this effect is reduced to being non-significant over time as the curves converge.

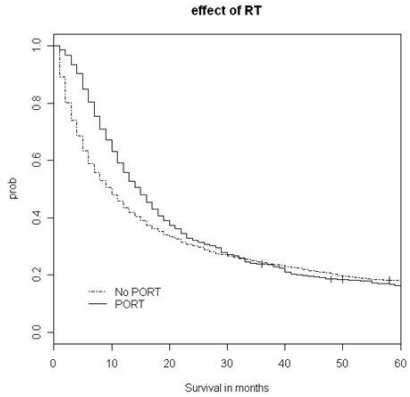

Figure 1. Effect of PORT on survival of gallbladder cancer.

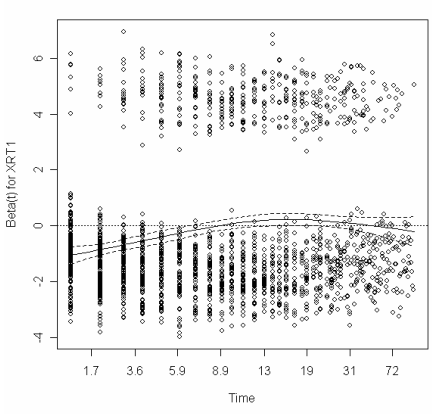

Figure 2. Time varying coefficient of PORT for gallbladder cancer.

Next, we created a RSF model for the data, which does not require proportionality of hazards. Interestingly, the C-index for the RSF model was 0.706, which was not statistically different from the C-index from the CPH model.

Subset analysis was performed to investigate the effect of PORT in the different populations by T classification and nodal status. The use of PORT for patients with node positive disease is associated with a significant increase in survival, as can be seen in Figure 3. However, we found that the use of PORT for node negative patients is associated with a decrease in survival.

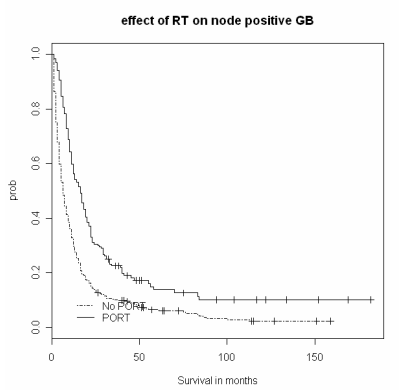

Figure 3. Effect of PORT on survival of node positive gallbladder cancer

Non-small cell lung cancer

The NSCLC dataset contained 23462 female and 27584 male patients. Of these, 8447 had received PORT while 42976 had not received PORT. 47911 patients had received a lobectomy while 3512 had a pneumonectomy. 21279 patients were coded as having a stage of T1, 24819 had a stage of T2, 1999 were T3 and 2949 were T4. 38385 patients were stage N0, 8421 were stage N1 and 4167 were stage N2.

As with the GB dataset, we first constructed a CPH model for the NSCLC dataset. The hazard ratios for the covariates in the CPH model are given in Table 3.

However, as we can see in Table 2, there are significant ($p < 0.05$) interactions between effect of PORT and the stage, especially for stages N2 and T3 and T4. The hazard ratios of these interaction terms are less than 1 indicating that PORT is associated with improved survival for these population subsets. PORT was associated with improved survival for N2 patients as seen in Figure 5.

	Hazard Ratio	Lower CI	Upper CI
Age	1.03	1.03	1.04
Age'	1.00	1.00	1.01
Age''	0.98	0.94	1.03
Sex=Male	1.29	1.26	1.32
Histology =Large cell CA	1.17	1.11	1.22
Histology =Other	0.83	0.79	0.87
Histology =Squamous Cell CA	1.04	1.02	1.07
Tstage=T2	1.28	1.24	1.31
Tstage=T3	2.05	1.90	2.20

Tstage=T4	1.78	1.68	1.90
Nstage=N 1	1.84	1.78	1.91
Nstage=N 2	2.65	2.51	2.80
XRT=RT	1.81	1.68	1.95
Site=pneu monectom y	1.22	1.17	1.27
Nstage=N 1 * XRT=RT	0.63	0.58	0.67
Nstage=N 2 * XRT=RT	0.50	0.46	0.54
Tstage=T3 * XRT=RT	0.68	0.61	0.77
Tstage=T4 * XRT=RT	0.82	0.74	0.92

Table 2. Hazards ratios for NSCLC dataset

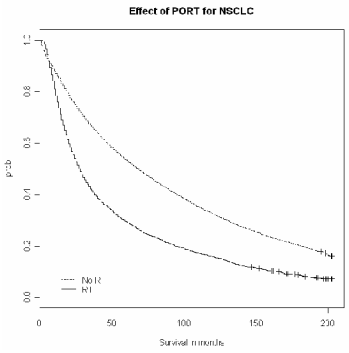

Figure 4. Effect of PORT on non-small-cell lung cancer

Overall, PORT had a detrimental effect on survival as seen in Figure 4. This was especially true for stages N0 and T1. This dataset is comprised largely of these groups and thus the effect on the overall survival is large.

Next, we constructed a RSF model with the NSCLC data. However, again the C-indices for the CPH and RSF models were not significantly different (p=0.66).

When we performed the hazard proportionality test for this dataset, there was significant variation over time, similar to the GB dataset, with PORT being one of the covariates with a time varying component.

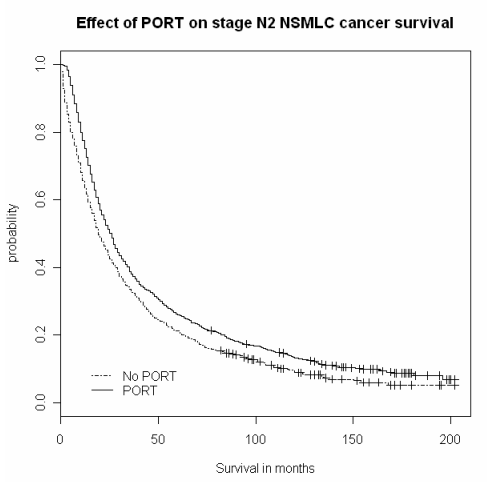

Figure 5. Effect of PORT on stage N2 non-small-cell lung cancer

We performed subset analyses to evaluate the interaction between PORT and the stage for both cancer sites. Table 4 summarizes the impact of PORT for the various stages. As we can see, there is a strong interaction between these covariates.

Site	Stage	Effect of PORT on survival	p-value
Gallbladder	T1	Reduced survival	0.0195
Gallbladder	T2	Increased survival	0.0168
Gallbladder	T3	Increased survival	1.65e-06
Gallbladder	T4	Increased survival	8.13e-08
Gallbladder	N1	Increased survival	9.99e-16
Lung	T1	Reduced survival	<0.0001
Lung	T2	Reduced survival	<0.0001
Lung	T3	no diff	0.849
Lung	T4	Reduced survival	2.90E-07
Lung	N0	Reduced survival	<0.00001
Lung	N1	no diff	0.0756
Lung	N2	Increased survival	2.60E-09
Lung	T1N2	Increased survival	1.20E-05
Lung	T2N2	Increased survival	0.00259
Lung	T3N2	Increased survival	0.0113
Lung	T4N2	Increased survival	0.00273

Table 3. Interaction of the effect stage and PORT on the survival of patients

Discussions

In general, both models performed reasonably well at making predictions for GB and NSCLC patients. These types of survival models can be used to help make more customized estimates of survival benefit for individual patients, which may be useful in the clinical setting when faced with the decision of making individualized recommendations of whether PORT would be worthwhile for a given patient. For example, for a 70 year old black male patient with T3N1 non-papillary GB cancer, the models predict that the addition of PORT would increase 3-year survival from 6% to 16.4%. For a 50 year old female with T4N1 non-papillary GB cancer, the models predict that adding PORT would increase 3-yr survival from 3.2 to 23.3%. Similarly, for a 50 year old white male with T1N2 adenocarcinoma NSCLC, the models predict that the addition of PORT would improve 5-yr survival from 45% to 49%, while the survival for a 70 year old white male with T4N2 large cell NSCLC would improve from 6.5% to 13.2%.

Initially, when we found that both of these datasets contained time-varying hazards, we hypothesized that RSF might perform better than CPH for these datasets, because RSF models do not require proportionality of hazards. When comparing the performance of these models, however, we found that the C-index did not vary significantly between these models. There may be several explanations for this. It may be because the influence of the specific covariates that had the time-varying component was small and thus did not have a significant effect on overall outcome. It may also be that there may be better validation measures, other than the C-index, that would demonstrate differences in performance. We are currently investigating alternative performance measures for these models.

There are several limitations to this study. The datasets used were retrospective data from the SEER database, and not from a randomized controlled clinical trial. The cases also span a long time period, and surgical techniques and radiation delivery techniques may have changed significantly over this time period. We only used one type of validation measure, the C-index, which is a measure of discrimination, but we did not evaluate calibration. Also, we only measured discrimination at one time interval (12 months). In the future, we may investigate the use of other validation measures, such as the Brier score.

Conclusion

Survival analysis can be used to assist clinicians and patients make informed personalized decisions about the benefit of post-operative radiation therapy. Our analysis found that PORT did provide benefit for node positive gallbladder patients, and for NSCLC patients with N2 disease. We found that although these datasets had time-varying covariates that violated the proportional hazard assumption, there was no significant difference in performance between the CPH model and the RSF models.

Acknowledgements

This research was funded under NLM training grant 2T15LM007088 and an Oregon Clinical Translational Research Institute pilot project grant.

References

1. Wang SJ, *et al,* Prediction model for estimating the survival benefit of adjuvant radiotherapy for gallbladder cancer, J Clin Oncol, 2008: 26.
2. Lally BE, *et al,* Postoperative Radiotherapy for Stage II or III non-small-cell lung cancer using the SEER database, J Clin Oncol, 2006, 24 (19):2998
3. Czito, BG., *et al,* Adjuvant external-beam radiotherapy with concurrent chemotherapy after resection of primary gallbladder carcinoma: A 23-year experience." International Journal of Radiation Oncology*Biology*Physics, 2005, 62(4) 103.
4. Stephens, R J, *et al* "The role of post-operative radiotherapy in non-small-cell lung cancer British Journal of Cancer 74, no. 4, 1996: 632-9.
5. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER 17 Regs Public-Use, Nov 2006 Sub (1973-2004), NCI, release April 2007.
6. Harrell F, Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer, New York, 2001.
7. Therneau T, A Package for Survival Analysis in S,http://mayoresearch.mayo.edu/mayo/research/biostat/upload/survival.pdf 1999
8. Breiman L, Random forests, Machine Learning, 45:5-32, 2001
9. Ishwaran H, *et al*, Random survival forests, Cleveland Clinic Technical Report, 2007
10. Hothorn T, *et al,* Survival ensembles, Biostatistics, 2006; 7(3):355-373