Predicting distant failure in early stage NSCLC treated with SBRT using clinical parameters

Z. Zhou[1], **M. Folkert**[1], **N. Cannon**[1], **P. Iyengar**[1], **K. Westover**[1], **Y. Zhang**, **H. Choy**[1], **R. Timmerman**[1], **J. Yan**[2], **X-J Xie**[2], **S. Jiang**[1], and **J. Wang**[1,*]

[1]Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, TX

[2]Department of Clinical Science, UT Southwestern Medical Center, Dallas, TX

Abstract

Purpose/Objective—The aim of this study is to predict early distant failure in early stage non-small cell lung cancer (NSCLC) treated with stereotactic body radiation therapy (SBRT) using clinical parameters by machine learning algorithms.

Materials/Methods—The dataset used in this work includes 81 early stage NSCLC patients with at least 6 months follow-up who underwent SBRT between 2006 and 2012 at a single institution. The clinical parameters (n=18) for each patient include demographic parameters, tumor characteristics, treatment fraction schemes, and pretreatment medications. Three predictive models were constructed based on different machine learning algorithms: 1) artificial neural network (ANN), 2) logistic regression (LR) and 3) support vector machine (SVM). Furthermore, to select an optimal clinical parameter set for the model construction, three strategies were adopted: 1) clonal selection algorithm (CSA) based selection strategy; 2) sequential forward selection (SFS) method; 3) statistical analysis (SA) based strategy. 5-cross-validation is used to validate the performance of each predictive model. The accuracy was assessed by area under the receiver operating characteristic (ROC) curve (AUC), sensitivity and specificity of the system was also evaluated.

Results—The AUCs for ANN, LR and SVM were 0.75, 0.73, and 0.80, respectively. The sensitivity values for ANN, LR and SVM were 71.2%, 72.9% and 83.1%, while the specificity values for ANN, LR and SVM were 59.1%, 63.6% and 63.6%, respectively. Meanwhile, the CSA based strategy outperformed SFS and SA in terms of AUC, sensitivity and specificity.

Conclusions—Based on clinical parameters, the SVM with the CSA optimal parameter set selection strategy achieves better performance than other strategies for predicting distant failure in lung SBRT patients.

*jing.wang@utsouthwestern.edu.

Conflicts of interest None.

1. Introduction

Dose escalation with stereotactic body radiation therapy (SBRT), using modern imaging and radiation delivery techniques, has improved local control outcomes in early stage non-small cell lung cancer (NSCLC) in patients who are not ideal surgical candidates [1, 2]. For medically inoperable patients, SBRT has been established as the standard of care for local control [3-5]. In a prospective trial study using 3-fraction SBRT for medically inoperable early stage NSCLC (RTOG 0236 trial) [6], excellent primary tumor local control rates of over 95% were observed after 3 years. Nevertheless, distant failure in early stage patients was still common, with a 3 year actuarial rate of 22.1%; notably, over 70% of these failures occurred within the first 2 years. In a recent abstract reporting an update on the RTOG 0236 experience, five-year distant failure rates of 31% were reported [7]. In a large retrospective series of 676 patients treated with SBRT for early stage NSCLC [8], the 2 and 5 year distant failure rates were 14.7% and 19.9%, respectively, with a median time to distant recurrence of 9.6 months. These data indicate that the vast majority of failures occur shortly after definitive treatment of the primary tumor, even in the setting of PET-based staging. Distant failure is a critical oncologic event, as it correlates very closely with mortality; patients with metastases outside the lung from NSCLC have a median survival of only 6 months [9]. For patients at high risk of early distant failure after SBRT treatment intensification with additional systematic therapy may reduce the risk of distant relapse and improve overall survival. A strategy that can correctly stratifies patients at high risk for failure is needed, as this population is generally in relatively poor health, and the toxicity of the therapy could itself contribute to increased mortality.

Here we designed machine learning based models [10-13] to predict distant failure in early stage NSCLC treated with SBRT using clinical parameters. While previous studies [7] [14] mainly focused on individual factors, the models developed in this work evaluated the possible inclusion of all available clinical parameters. Specifically, support vector machine (SVM) [15, 16] model, artificial neural network (ANN) [17] and logistic regression (LR) [18] based models were adopted for constructing the predictive model. These three models were chosen because they were the most common used methods to construct the predictive models [19-22]. Furthermore, to select optimal clinical parameters and train model parameters, we investigated three strategies including:1) clonal selection algorithm (CSA) [23] based method; 2) sequential forward selection (SFS) method; 3) statistical analysis (SA) based method.

2. Materials and Methods

2.1 Patients and clinical parameters

The retrospective study was approved by the institutional review board (IRB). The cohort included 81 early stage (Stage IA and IB) NSCLC patients. These patients underwent SBRT

from 2006 to 2012 at our institute with at least 6 months of follow-up. The range of follow-up was from 6 months to 74 months and the median follow-up time was about 18 months. These patients were generally followed up at 8 weeks post treatment, then every 3 months for 2 years, and then every 6 months for 2 years, and then annually thereafter. Twenty three patients (27.2%) failed at distant sites. Patients underwent regular imaging of the chest, as well as directed imaging based on any clinical findings, which could include imaging of the abdomen and pelvis (generally CT-based) and CT or MRI for brain and spine (MRI preferred unless otherwise contraindicated). PET scans were generally performed if the criteria for local enlargement on CT was present (at least 20% increase in the longest diameter) to help distinguish recurrence from post-radiation changes, or for restaging due to clinical or radiographic evidence of progressive disease. The median interval between the distant failure and last SBRT treatment fraction was 10 months. Demographic features and clinical parameters for each patient were extracted from clinical charts.

The clinical parameters for each patient were categorized into four groups: (1) demographic parameters (Table S1.A); (2) tumor characteristics (Table S1.B); (3) treatment parameters (Table S1.C); and (4) pretreatment medications (Table S1.D). Each of these clinical parameters was used as an independent feature for the predictive model.

2.2. Predictive model construction

Each clinical parameter was utilized as an independent input characteristic for the optimal feature (i.e., clinical parameter) set selection and model construction. Cross-validation was employed to remove bias in selecting features and avoid over fitting during constructing the prediction model [19]. As there was a large difference in patient numbers between the two groups (23 patients with distant failure and 58 patients without distant failure), an additional strategy was needed to overcome the influence of imbalanced data. In this work, an over-sampling technique called synthetic minority over-sampling technique (SMOTE) [24] was applied. In SMOTE, the minority class was oversampled by using K-nearest neighborhood (KNN) graph. The SMOTE-generated new samples were then added to the training set for constructing the predictive model.

2.3 Feature selection

In this work, three feature selection methods including clonal selection algorithm (CSA), sequential forward feature selection (SFS) and statistical analysis based method (SA) were utilized. SA is the most common used method in medical analysis [25], while SFS is one of the most classical methods for feature selection in machine learning approaches[26]. CSA is a recently developed method which may improve the performance of prediction accuracy[27]. Therefore, they were utilized in this study and the details of these feature selection methods were described in the Supplement.

3. Results

Figure 1 shows the AUC, sensitivity and specificity for three models with the CSA feature selection strategy. The AUCs for SVM, ANN and LR models were 0.80, 0.75 and 0.73, respectively. The sensitivities for SVM, ANN and LR were 83.1%, 71.2% and 72.9%, while

the specificities for SVM, ANN and LR were 63.6%, 59.1% and 63.6%, respectively. Table 1 shows the selected optimal clinical parameters for each of three predictive models. The number of selected clinical parameters for SVM, ANN and LR were 6, 12, and 8, respectively. Data show that SVM selects fewer clinical parameters than the other two models. Figure 2 presents the corresponding ROC curves for the three models. It was observed that the SVM achieved the best performance among three predictive models investigated in this work. In addition, by analyzing the odds ratio, the effects of the selected optimal clinical parameters in three models were shown in Table S2 in the Supplement. Parameters associated with one or zero distant failure were not compared.

Table 2 presents the AUC, sensitivity and specificity for three strategies using the SVM predictive model. CSA outperformed SFS and SA in all the three evaluation criteria. Fig. 3 shows the ROC curves for three methods. AUCs were 0.80, 0.75 and 0.61, respectively. The gain of CSA is further demonstrated by the AUC measure. It was shown that CSA is significantly better than SFS and SA ($P<0.0001$) according to the unpaired t test with a 95% confidence interval.

4. Discussion

Non-small cell lung cancer is an aggressive disease, but significant improvements in survival can be achieved if found and treated early. Based on the National Lung Screening Trial (NLST), identification of disease at an early stage allowed for prompt treatment, resulting in a relative reduction of lung cancer mortality by 20% [28]. Data from prospective trials showed that even early stage T1-2N0 lung lesions receiving optimal treatment with well-controlled primary lesions will develop distant metastases, ultimately leading to mortality. The addition of adjuvant systemic therapy may help reduce this risk, but would result in overtreatment of 69-80% of patients, at a significant financial cost to the medical care system as well as a human cost to patients, given that chemotherapy also carries a significant toxicity risk (including death from infection secondary to myelosuppression) as well as negative impact on patient quality of life. The identification of patients at high risk for early recurrence is imperative to guide the application of additional therapy to the population that would most likely benefit from it, and such prognostic models are potentially the backbone for risk-adapted prospective therapeutic protocols. Patients at risk for early distant failure are the most likely to benefit from an early systemic intervention, as they have just gone through a thorough staging process and any metastatic disease would be as small as possible, below the limits of detection of standard imaging.

In this work, we built a robust predictive model for early distant failure following definitive SBRT for early stage NSCLC. We investigated the performance of three different machine learning based classifiers. Our results show that the SVM model outperforms both ANN and LR in terms of AUC, sensitivity and specificity. Using clinical parameters and demographic features, the SVM model can achieve 0.80 AUC. The performance of predictive models could be further improved by incorporating additional input features, such as imaging characteristics of pre- and post-SBRT positron emission tomography (PET) and dynamic contrast-enhanced MRI.

In addition to the predictive models, feature selection strategy is another important factor affecting the performance of machine learning based algorithms. Three strategies were investigated in this work. Compared to SFS and SA, CSA can improve prediction accuracy (measured by AUC), sensitivity, and specificity. Especially for the SVM model, CSA not only selected an optimal clinical parameter set, but also obtained optimal model parameters for SVM. It was also noted that different optimal feature sets are selected for different prediction models. Nevertheless, stage and location were selected in all the three models while tumor size, ethnicity, histology, anti-inflammatory drugs, ACE inhibitors and ASA were selected in two out three models.

We investigated three machine learning based models for distant failure prediction of early stage NSCLC treated with SBRT. We also studied three feature selection strategies to obtain an optimal clinical parameter set for each of the prediction models. By selecting clinical parameters and training SVM parameters based on CSA, the SVM model achieved higher prediction accuracy than ANN and LR models. Meanwhile, we observed that CSA based feature selection strategy can obtain better performance than SFS and SA. We will seek to validate these models with a similar cohort of patients from a collaborating institution once a research agreement has been established.

The ultimate goal of this model development is to implement/integrate a successful model into a prospective trial for treatment intensification in patients with early stage NSCLC treated with SBRT. After the treatment, the patients would be stratified to receive adjuvant chemotherapy or observation. Despite significant underlying clinical differences, we could also investigate implementation of this type of prognostic model for early stage NSCLC treated with surgery, as similar distant failure patterns exist in this population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

References

[1]. Høyer M. Improved accuracy and outcome in radiotherapy of lung cancer. Radiotherapy and Oncology. 2008; 87:1–2. [PubMed: 18378029]

[2]. van Baardwijk A, Tomé WA, van Elmpt W, Bentzen SM, Reymen B, Wanders R, et al. Is high-dose stereotactic body radiotherapy (SBRT) for stage I non-small cell lung cancer (NSCLC) overkill? A systematic review. Radiotherapy and Oncology. 2012; 105:145–9. [PubMed: 23068707]

[3]. Chi A, Liao Z, Nguyen NP, Xu J, Stea B, Komaki R. Systemic review of the patterns of failure following stereotactic body radiation therapy in early-stage non-small-cell lung cancer: clinical implications. Radiotherapy and Oncology. 2010; 94:1–11. [PubMed: 20074823]

[4]. Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, et al. Non–Small Cell Lung Cancer, Version 6.2015. Journal of the National Comprehensive Cancer Network. 2015; 13:515–24. [PubMed: 25964637]

[5]. Liu H-W, Gabos Z, Ghosh S, Roberts B, Lau H, Kerba M. Outcomes in stage I non-small cell lung cancer following the introduction of stereotactic body radiotherapy in Alberta–A population-based study. Radiotherapy and Oncology. 2015; 117:71–6. [PubMed: 26349589]

[6]. Timmerman R, Paulus R, Galvin J, Michalski J, Straube W, Bradley J, et al. Stereotactic Body Radiation Therapy for Inoperable Early Stage Lung Cancer. Jama-J Am Med Assoc. 2010; 303:1070–6.

[7]. Timmerman RD, Hu C, Michalski J, Straube W, Galvin J, Johnstone D, et al. Long-term Results of RTOG 0236: A Phase II Trial of Stereotactic Body Radiation Therapy (SBRT) in the Treatment of Patients with Medically Inoperable Stage I Non-Small Cell Lung Cancer. Int J Radiat Oncol. 2014; 90:S30–S.

[8]. Senthi S, Lagerwaard FJ, Haasbeek CJ, Slotman BJ, Senan S. Patterns of disease recurrence after stereotactic ablative radiotherapy for early stage non-small-cell lung cancer: a retrospective analysis. The lancet oncology. 2012; 13:802–9. [PubMed: 22727222]

[9]. Postmus PE, Brambilla E, Chansky K, Crowley J, Goldstraw P, Patz EF Jr, et al. The IASLC Lung Cancer Staging Project: proposals for revision of the M descriptors in the forthcoming (seventh) edition of the TNM classification of lung cancer. Journal of Thoracic Oncology. 2007; 2:686–93. [PubMed: 17762334]

[10]. Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CM, et al. Rapid Learning health care in oncology'–An approach towards decision support systems enabling customised radiotherapy. Radiotherapy and Oncology. 2013; 109:159–64. [PubMed: 23993399]

[11]. van Stiphout RG, Lammering G, Buijsen J, Janssen MH, Gambacorta MA, Slagmolen P, et al. Development and external validation of a predictive model for pathological complete response of rectal cancer patients including sequential PET-CT imaging. Radiotherapy and Oncology. 2011; 98:126–33. [PubMed: 21176986]

[12]. Roelofs E, Dekker A, Meldolesi E, van Stiphout RG, Valentini V, Lambin P. International data-sharing for radiotherapy research: an open-source based infrastructure for multicentric clinical data mining. Radiotherapy and Oncology. 2014; 110:370–4. [PubMed: 24309199]

[13]. Kazmierska J, Malicki J. Application of the Naïve Bayesian Classifier to optimize treatment decisions. Radiotherapy and Oncology. 2008; 86:211–6. [PubMed: 18022719]

[14]. Clarke K, Taremi M, Dahele M, Freeman M, Fung S, Franks K, et al. Stereotactic body radiotherapy (SBRT) for non-small cell lung cancer (NSCLC): is FDG-PET a predictor of outcome? Radiotherapy and Oncology. 2012; 104:62–6. [PubMed: 22682749]

[15]. Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995; 20:273–97.

[16]. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST). 2011; 2:27.

[17]. Rochester N, Holland J, Haibt L, Duda W. Tests on a cell assembly theory of the action of the brain, using a large digital computer. Information Theory, IRE Transactions on. 1956; 2:80–93.

[18]. Freedman, DA. Statistical models: theory and practice. cambridge university press; 2009.

[19]. Zhang H, Tan S, Chen W, Kligerman S, Kim G, D'Souza WD, et al. Modeling Pathologic Response of Esophageal Cancer to Chemoradiation Therapy Using Spatial-Temporal 18 F-FDG PET Features, Clinical Parameters, and Demographics. International Journal of Radiation Oncology* Biology* Physics. 2014; 88:195–203.

[20]. Zhou Z-G, Liu F, Jiao L-C, Wang Z-L, Zhang X-P, Wang X-D, et al. An evidential reasoning based model for diagnosis of lymph node metastasis in gastric cancer. BMC medical informatics and decision making. 2013; 13:123. [PubMed: 24195733]

[21]. Zhang X-P, Wang Z-L, Tang L, Sun Y-S, Cao K, Gao Y. Support vector machine model for diagnosis of lymph node metastasis in gastric cancer with multidetector computed tomography: a preliminary study. BMC cancer. 2011; 11:1.

[22]. Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nature medicine. 2001; 7:673–9.

[23]. De Castro LN, Von Zuben FJ. Learning and optimization using the clonal selection principle. Evolutionary Computation, IEEE Transactions on. 2002; 6:239–51.

[24]. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res. 2002; 16:321–57.

[25]. Breslow, NE.; Day, NE. Statistical methods in cancer research. International Agency for Research on Cancer Lyon; 1987.

[26]. Jain A, Zongker D. Feature selection: Evaluation, application, and small sample performance. Ieee T Pattern Anal. 1997; 19:153–8.

[27]. Zhang L, Zhong Y, Huang B, Gong J, Li P. Dimensionality reduction based on clonal selection for hyperspectral imagery. Geoscience and Remote Sensing, IEEE Transactions on. 2007; 45:4172–86.

[28]. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England journal of medicine. 2011; 365:395–409. [PubMed: 21714641]
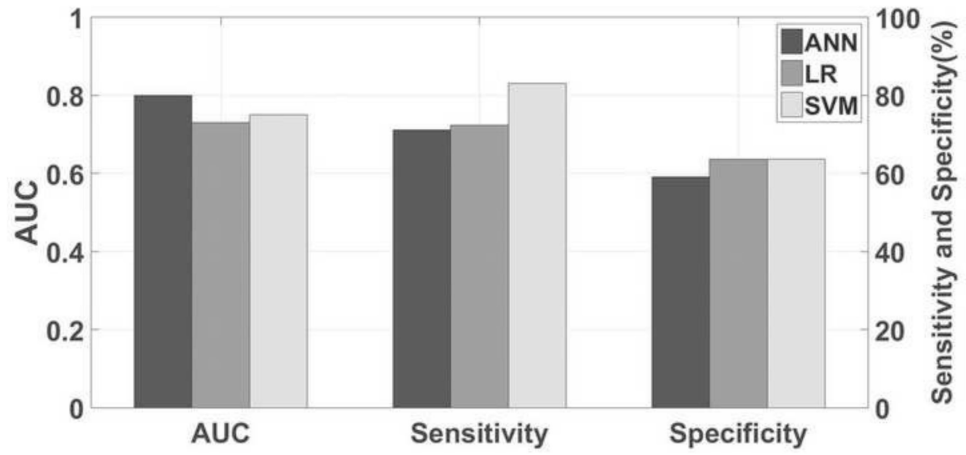
Fig. 1.
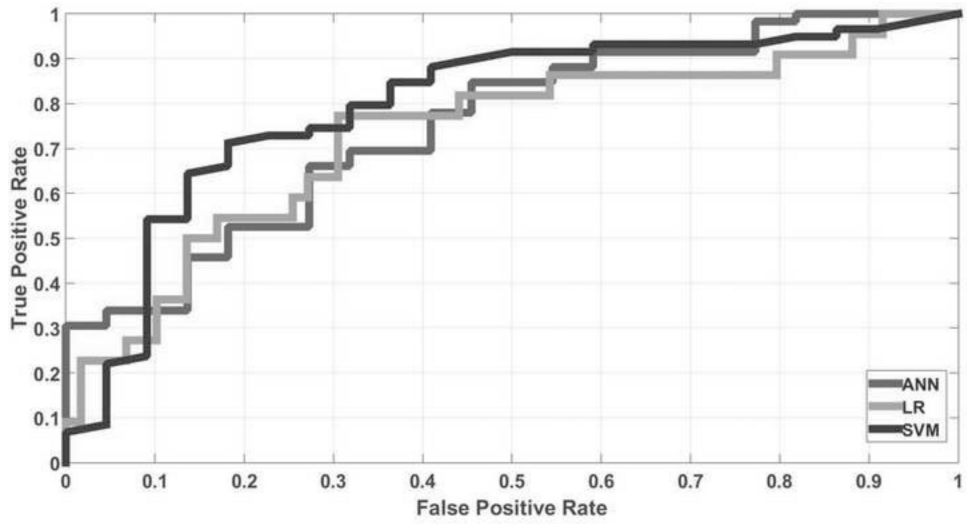AUC, sensitivity and specificity for three predictive models; (B) ROC curves for three predictive models.
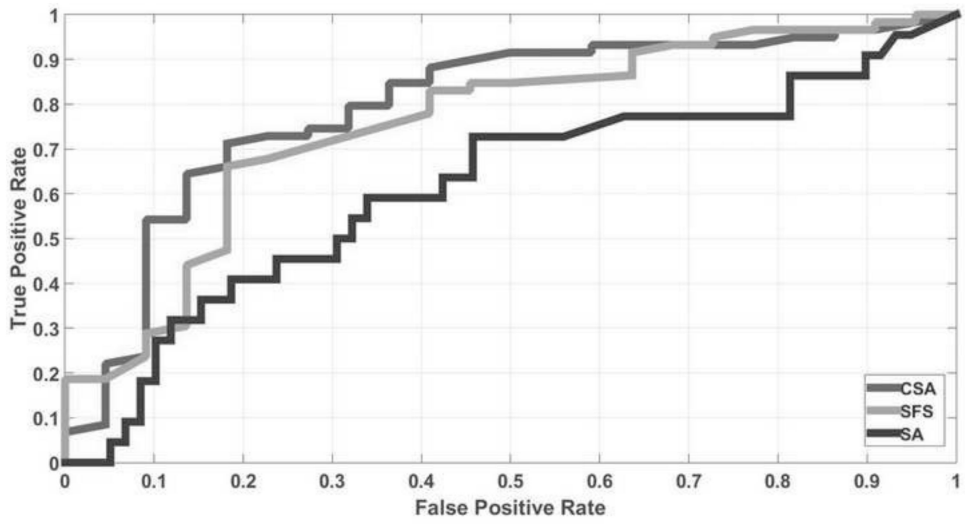
Fig. 2.
ROC curves for three predictive models.

Fig. 3.
ROC curves for CSA, SFS and SA feature selection strategy in the SVM predictive model.

Table 1

Selected clinical parameters of three predictive models.

ANN	LR	SVM
Age	Ethnicity	Tumor size
Ethnicity	Histology	Location
Central tumor or not	Location	Stage
Tumor size	Stage	Metformin
Histology	Number fraction	Statin
Location	Dose per fraction	ACEinhibitor
Stage	Antiinflammatories	
Antiinflammatories	ASA	
Anitdiabetic		
Metformin		
ACEinhibitor		
ASA		

Table 2

AUC, sensitivity and Specificity of three feature selection methods for the SVM predictive model.

	AUC	Sensitivity (%)	Specificity (%)
CSA	0.80	83.1	63.6
SFS	0.75	77.9	59.1
SA	0.61	72.8	45.5