

Published in final edited form as:

Phys Med Biol. 2008 January 7; 53(1): 203–216. doi:10.1088/0031-9155/53/1/014.

Using patient data similarities to predict radiation pneumonitis via a self-organizing map

Shifeng Chen, Sumin Zhou, Fang-Fang Yin, Lawrence B Marks, and Shiva K Das

Department of Radiation Oncology, Duke University Medical Center, Durham, NC 27710, USA

Abstract

This work investigates the use of the self-organizing map (SOM) technique for predicting lung radiation pneumonitis (RP) risk. SOM is an effective method for projecting and visualizing high-dimensional data in a low-dimensional space (map). By projecting patients with similar data (dose and non-dose factors) onto the same region of the map, commonalities in their outcomes can be visualized and categorized. Once built, the SOM may be used to predict pneumonitis risk by identifying the region of the map that is most similar to a patient's characteristics. Two SOM models were developed from a database of 219 lung cancer patients treated with radiation therapy (34 clinically diagnosed with Grade 2+ pneumonitis). The models were: SOM_{all} built from all dose and non-dose factors and, for comparison, SOM_{dose} built from dose factors alone. Both models were tested using ten-fold cross validation and Receiver Operating Characteristics (ROC) analysis. Models SOM_{all} and SOM_{dose} yielded ten-fold cross-validated ROC areas of 0.73 (sensitivity/specificity = 71%/68%) and 0.67 (sensitivity/specificity = 63%/66%), respectively. The significant difference between the cross-validated ROC areas of these two models ($p < 0.05$) implies that non-dose features add important information toward predicting RP risk. Among the input features selected by model SOM_{all}, the two with highest impact for increasing RP risk were: (a) higher mean lung dose and (b) chemotherapy prior to radiation therapy. The SOM model developed here may not be extrapolated to treatment techniques outside that used in our database, such as several-field lung intensity modulated radiation therapy or gated radiation therapy.

1. Introduction

Lung radiation pneumonitis (RP) is one of the major toxicities experienced by cancer patients receiving thoracic radiation therapy (RT) (Kocak *et al* 2005a). It occurs in approximately 13–37% of patients who undergo radiotherapy for lung cancer (Rodrigues *et al* 2004). Identification of factors that predict for the incidence of RP is important for reducing its probability of occurrence. For example, should adjuvant chemotherapy be identified as a risk factor, then alteration of chemotherapy regimens can be considered to reduce RP incidence.

Various predictors have been identified to predict the risk of RP (Rodrigues *et al* 2004). Three such predictors, derived from the lung dose–volume histogram, are frequently reported to be significantly correlated with RP: V_x , the volume above x Gy, for example V_{20} (Kong *et al* 2006, Tsujino *et al* 2006, 2003, Chang *et al* 2006, Rancatiet *al* 2003, Jenkinset *al* 2003, Moiseenko *et al* 2003) and V_{15} (Tsujino *et al* 2006, Schallenkamp *et al* 2007), mean lung dose (MLD) (Kong *et al* 2006, Chang *et al* 2006, Graham *et al* 1999, Jenkins *et al* 2003, Hernando *et al* 2001, Martel *et al* 1994, Kwa *et al* 1998, Theuws *et al* 1998b) and

the Lyman (1985) normal tissue complication probability (NTCP) (Kong *et al* 2006, Tsujino *et al* 2006, Hernando *et al* 2001, Seppenwoolde *et al* 2003, Marks *et al* 1997). These predictors, while important for understanding the causes of RP, generally only have poor-to-fair predictive ability (Rodrigues *et al* 2004).

Dose factors alone may not be ideal predictors (Marks 2002). Non-dose factors such as age, gender, tumor location, chemotherapy, etc, although nonsignificant by themselves, could likely contribute significantly when combined with other dose and non-dose factors in a multivariate model. This synergistic enhancement of predictive capability is suggested by the findings in Marks *et al* (1997), who reported that the correlation of Lyman NTCP and V_{30} with RP improved when patients with the poorest pre-RT pulmonary function were excluded. In a similar vein, Lind *et al* (2006) reported that the correlation of V_{20} with RP was greater in patients younger than 55 years of age. This appears to warrant a predictive model that is somehow capable of discerning and quantifying these interactions. For example, Das *et al* (2007) developed a decision tree model to predict the risk of RP, and demonstrated that the model including non-dose factors is superior to that using dose factors alone. The aim of this work is to develop and test an RP predictor based on self-organizing maps (SOM) (Kohonen 1995), a technique that is potentially capable of extracting and quantifying the underlying synergistic interactions.

SOM is a type of artificial neural network that is trained using unsupervised learning, capable of projecting and visualizing high-dimensional data in a low-dimensional space (Kohonen 1995). The SOM technique clusters and classifies patients into neurons based on similarities in the patient input data, unlike other techniques which additionally take patient outcomes into consideration. By searching for data similarities, the SOM is capable of extracting synergistic interactions between factors that are more powerfully predictive than the individual factors (Guyon and Elisseeff 2003). SOM has been previously applied in medicine (e.g. computer-aided detection of breast cancer (Markey *et al* 2003, Chen *et al* 2000)). To the best of our knowledge, SOM has never been applied to predict the risk of radiation pneumonitis.

In this work, the SOM model was built with input features selected from the total available set of factors, using a unique methodology. Predictive capability of the model was realistically tested using ten-fold cross validation (Hastie *et al* 2002). The effect of non-dose variables on predictive capability was assessed by comparing the SOM model built from dose alone to that including non-dose factors. Model comparison used receiver operating characteristic (ROC) curves (Das *et al* 2005, Lind *et al* 2002, Swets and Pickett 1982). Lastly, the importance of individual input features in the SOM model was evaluated.

2. Materials and methods

2.1. Input data

The patient database used in this study was previously described by Kocak *et al* (2005b). The database consists of 235 patients with lung cancer who received three-dimensional conformal, external-beam RT at Duke University Medical Center on an Institutional Review Board approved protocol (protocol title: prospective study of the influence of external beam radiation on pulmonary function (whole organ and regional) with radiographic correlations; protocol number: IRB 1698). Table 1 lists the patient and treatment characteristics. The patient population in this database was treated with RT delivered via parallel opposed anterior-posterior fields, followed by off-cord oblique fields. Most patients (70%) were treated at 1.8–2.0 Gy/fraction, once daily. The remaining were treated at 1.25 Gy/fraction to the clinical target volume and 1.6 Gy/fraction to the gross volume, twice daily (minimum interval was 6 h between two fractions). Dose distributions were three-dimensionally

computed based on computed tomography (CT) scans of patients in the treatment position. The left and right lungs were considered as a paired organ (excluding planning target volume).

Radiation-induced symptomatic pneumonitis was assessed at 1, 3 and then every 3–4 months post-RT. Based on the modified National Cancer Institute Common Toxicity Criteria (NCICTC) (table 2), pneumonitis was graded from 0 to 4. These NCICTC criteria are similar to the Southwestern Oncology Group (SWOG) toxicity criteria (Green and Weiss 1992) (table 2), often used in other studies (Kwa *et al* 1998, Seppenwoolde *et al* 2003). The toxicity endpoint of this study, Grade 2+ pneumonitis, is equivalent in both sets of criteria. Of the 235 patients in this analysis, 37 and 13 patients experienced Grade 2 and Grade 3 RP, respectively, and none experienced Grade 4 RP. Among patients with Grade 2 RP, 16 patients were classified as ‘hard-to-score’ due to uncertain diagnosis (Kocak *et al* 2005a). For the purpose of this study, 34 patients (excluding hard-to-score patients) were considered as having contracted Grade 2+ pneumonitis.

A large number of factors were extracted from the patient database, explained next. This large set comprised all factors that we believed might have potential to be selected as factors by the SOM. No preliminary prospective analysis was employed to include/exclude factors, since such an analysis may not necessarily identify factors that could be selected by the SOM as strongly synergistic with other factor(s) in the database.

A total of 27 non-dose factors (biological, clinical and other factors) were collected for each patient. These factors included race, age, gender, tumor stage, tumor location (central (mediastinal) or peripheral; upper, middle or lower lobe; right or left lung), chemotherapy schedule (none, pre-RT, concurrent, pre-RT and concurrent, post-RT or concurrent and post-RT), histology type (squamous cell, adenocarcinoma, non-small cell, small cell, large cell or other), surgery (yes versus no), once or twice daily RT, pre-RT FEV₁ (forced expiratory volume in 1 s), FEV₁% (as percentage of predicted normal), pre-RT DLCO (carbon monoxide diffusion capacity in lung) and pre-RTDLCO% (as percentage of predicted normal). FEV₁ and DLCO% were not collected for approximately 20% of patients. In accordance with statistical treatment of missing values (Hastie *et al* 2002, Setiono 2001), they were assigned the average of non-missing values.

The lung cumulative dose–volume histogram (DVH) was calculated for each patient, for a bin size of 2 Gy. The lung DVH was also reduced to the generalized equivalent uniform dose (EUD) (Niemierko 1999) with the equation

$$\text{EUD} = \left(\frac{\sum_i v_i D_i^a}{\sum_i v_i} \right)^{1/a}, \quad (1)$$

where v_i is the lung volume receiving dose D_i . EUDs were generated for the exponential parameter a ranging from 0.4 to 4, in increments of 0.1. Note that for $a = 1$, EUD is equivalent to MLD. In addition to the lung DVH and lung EUDs, the biological mean lung dose (NTD_{mean}) was also included as a factor. NTD_{mean} was calculated by normalizing the lung DVH to a fraction size of 2 Gy (Lebesque and Keus 1991):

$$\text{NTD}_i = \frac{nd_i \{1 + d_i / (\alpha/\beta)\}}{1 + 2 / (\alpha/\beta)}, \quad (2)$$

where n is the number of fractions, d_i is the dose per fraction for lung volume v_i and α/β is the linear quadratic model parameter (assumed to be $\alpha/\beta = 3$ Gy (Dubray *et al* 1995, Kwa *et al* 1998) in this study). Similar to the physical mean lung dose, the biological mean lung dose, NTD_{mean} , was calculated as

$$NTD_{\text{mean}} = \frac{\sum_i v_i NTD_i}{\sum_i v_i}. \quad (3)$$

Thus, the lung dose factors considered in this analysis were the DVH points, V_x (volume above x Gy), for x ranging from 6 Gy to 60 Gy in increments of 2 Gy, EUDs for the exponents $a = 0.4$ to 4 (increments of 0.1) and NTD_{mean} .

Luijk *et al* (2005) reported that the tolerance dose for early lung function damage also depends on the concomitant irradiation of the heart. Consequently, we included the mean dose received by the heart as a factor in this analysis.

2.2. Data partitioning for model creation and testing: cross validation

For small datasets, K -fold cross validation is the recommended method to assess the performance of a prediction model (Hastie *et al* 2002). In this technique, the data are randomly split into K approximately equal-sized groups. $K - 1$ groups (training data) are used to create (train) the model and the remaining group (cross-validation data) is used as a test to measure the performance of the prediction model. The procedure is repeated K times, with each group, in turn, serving as the cross-validation data. The typical choices of K are 5 or 10 (Hastie *et al* 2002). In this analysis, we chose ten-fold cross validation ($K = 10$).

2.3. Model creation and testing: self-organizing map

The basic idea behind the SOM is to construct a nonlinear transformation to project high-dimensional input data onto a low, often two-dimensional space, termed a feature map. This makes it particularly useful for patient data, i.e. since it is difficult to search for similarities in high dimensions, SOM provides a ‘collapsed’ view that is easier to interpret. The disadvantage of SOM is that this reduction from high to low dimensions might also inadvertently conceal similarities that exist in high dimensions. In this work, similar patient input vectors in the higher dimensional input space were clustered into the same region or neuron of the feature map. The entries in the input vector, termed features, are a subset of all available factors. The algorithm for selection of input features is detailed later in this section. For example, if the input features constituting the input vector are ‘ a ’ and ‘ b ’, patients with similar values of a and b were clustered into the same region (neuron). Each neuron in the map contains two parts of information. The first part is its relative physical location, i.e. proximity to other neurons, and the second part is its reference vector or weight vector, which can be thought of as the ‘typical’ feature values for that neuron.

During the course of SOM training, patients were assigned to neurons with reference vectors most similar to the patient input vector. Following each such patient assignment, the reference vector of the assigned neuron and its immediate neighboring neurons were updated to reflect the change to the patient population at the neuron. The neighboring neurons were identified as neurons whose distances to the assigned neuron were less than the neighborhood distance (nd). The neighborhood distance (nd) decreased from a large number to 0 during training, such that the neighborhood neurons included all neurons at the beginning and only the assigned neuron at the end. The learning rate (lr) determined the amount of information passed from the patient input vector to the assigned neuron, which decreased from 1 to 0 during training. Similarity was judged by the Euclidean distance

between a patient input vector and the neuron reference vector (a smaller distance implies greater similarity). Since SOM is an unsupervised learning method, patients were assigned to each neuron without knowledge of their toxicity classification (RP versus no RP). The training algorithm for SOM with M neurons is briefly summarized by the following steps:

1. Initialize the reference vector of each neuron to random values, and set the initial neighborhood distance (nd) and initial learning rate (lr).
2. for $i = 1$ to N (N is the number of patients used to train SOM)
 - a. Randomly select one unused patient (input vector \mathbf{x}).
 - b. for $j = 1$ to M , calculate the Euclidean distance (d_j) between the patient and the reference vector for neuron j (\mathbf{w}_j), $d_j = \|\mathbf{x} - \mathbf{w}_j\|$.
 - c. Determine the best matching neuron o with the minimal distance, $d_o = \min_{j=1:M} (d_j)$.
 - d. for $j = 1$ to M , update the reference vector of the neuron j : $\mathbf{w}_j \rightarrow \mathbf{w}_j + lr \cdot f(d_{oj}) \cdot (\mathbf{x} - \mathbf{w}_j)$

$$f(d_{oj}) = \begin{cases} 1, & d_{oj} = 0 \\ 0.5, & 0 < d_{oj} \leq nd \\ 0, & d_{oj} > nd, \end{cases}$$

where d_{oj} is the Euclidean distance between neuron j and the best matching neuron o .

- e. Decrease the neighborhood distance (nd) and the learning rate (lr): $nd = nd - \Delta nd$, and $lr = lr - \Delta lr$.
3. Repeat step 2 until convergence.

The default values of SOM toolbox in MATLAB (Math Works, Inc., Natick, MA) were used for the initial neighborhood distance (nd), the initial learning rate (lr), step Δnd and step Δlr .

Input features were selected from the list of 93 dose and non-dose factors by a trial addition/substitution process, as follows. In the ten-fold cross-validation scheme described in the previous subsection, nine groups of data were used to train the SOM. To optimally select the input features, the SOM was trained (MATLAB, The Math Works, Inc., Natick, MA) using eight of the nine training groups and then evaluated on the remaining training group by ROC analysis (greater model accuracy implies a higher area under the ROC curve, which plots sensitivity versus 1-specificity for varying values of the threshold separating cases with and without RP (Swets and Pickett 1982)). The eight training groups used to build the SOM were collectively termed the training-construction set, and the one training group used for evaluation was termed the training-evaluation set. This evaluation process was internal to the training group and only served to build the list of selected features, i.e. it was not used as an unbiased test of SOM performance. The added or substituted factor was accepted as an input feature if the area under the training-evaluation ROC curve increased. SOM construction was stopped if no new factor was accepted as an input feature. To prevent the selection of multiple highly correlated dose factors as features, the SOM disallowed selection of any dose factor having correlation >0.95 with any of the already selected features.

Once the SOM was trained, the probability of RP at every neuron was computed as

$$P = \frac{N_p}{N_n + N_p}, \quad (4)$$

where N_n and N_p are the numbers of patients suffering RP and not suffering RP, respectively, at that neuron. For prospective use (i.e. to predict the risk of RP for patients outside the database used to build the SOM), the SOM-predicted risk of a patient suffering RP is the probability associated with the closest matched neuron (neuron most resembling the patient).

Unbiased testing of the SOM followed construction, wherein each test group was evaluated using the SOM built from the training set. Since each of the ten patient groups sequentially served as the test group, ten SOMs were built for the purpose of unbiased testing. The combined results of all ten test groups were then analyzed to evaluate SOM predictive ability using the metrics: area under the ROC curve (AUC), sensitivity and specificity.

2.4. Model comparison and evaluation of importance of input features

To evaluate the influence of non-dose factors in modeling the risk of RP, two SOM models were built: SOM_{dose} and SOM_{all} , with input features selected from dose factors alone and all factors, respectively. ROCKIT software (Metz *et al* 1998) from the Department of Radiology, University of Chicago, was used to evaluate the statistical significance of the difference in areas between the ROC curves from the two models. Statistical significance was based on a univariate z -score test (null hypothesis: the datasets arose from binormal ROC curves with equal areas beneath them).

The importance of input features was evaluated in the SOM_{all} model by individual exclusion. Each time a feature was excluded from SOM_{all} , the model was retrained and tested. The cross-validated ROC area decrement resulting from feature exclusion was used to rank its importance: a larger decrement signifies a more important feature. Statistical significance of the feature was judged by statistical significance of the area decrement (ROCKIT (Metz *et al* 1998)).

3. Results and discussion

3.1. Model building

The optimal SOM configuration was chosen as a 4×3 neuron map after evaluating various configurations (see table 3). The AUC for training data continuously increases with increasing neurons, whereas the AUC for training-evaluation data decreases beyond 12 neurons (the 4×3 map). Training evaluation provides an effective way to avoid model under-training (model is not complex enough) or over-training (model is too complex—it fits the signal as well as the noise).

The input features selected by models SOM_{all} and SOM_{dose} are shown in table 4. The table displays the input features selected by all ten of the SOMs used for cross validation (one SOM for each test set). Since the training data for each SOM is slightly different (two of nine groups are different for any two SOMs), the selected input features vary slightly between the SOMs. Statistical significance of the features selected by model SOM_{all} is discussed in the next subsection.

For model SOM_{all} , the ten cross-validation SOMs selected three equivalent uniform doses (EUD) as input features, with exponents $a = 0.9, 1.0$ and 1.1 . These three dose factors, while highly correlated to each other, were selected by different cross-validation SOMs, i.e. no

two of these dose features were selected by the same cross-validation SOM. Since EUD $a = 1$ corresponds to MLD, the dose features selected by SOM_{all} can be approximately characterized as MLD. MLD frequently appears as a strong predictor of RP risk in the literature (Kong *et al* 2006, Chang *et al* 2006, Graham *et al* 1999, Jenkins *et al* 2003, Hernando *et al* 2001, Martel *et al* 1994, Kwa *et al* 1998, Theuws *et al* 1998b). For model SOM_{dose}, the ten cross-validated SOMs also selected EUDs similar to the MLD (exponents $a = 0.7, 0.8, 0.9$ and 1.0). Additionally, SOM_{dose} also selected three V_x values ($x = 40, 42, 44$ Gy). Again, no two EUD values or two V_x values with correlation >0.95 were selected by the same cross-validated SOM. (Note that the selected features did not change for dose bin sizes <2 Gy.) In the literature, V_x for $x = 20, 30, 15$ Gy, lower than those selected by SOM_{dose}, are reported as predictive for RP (Kong *et al* 2006, Tsujino *et al* 2006, Lind *et al* 2006, Chang *et al* 2006, Graham *et al* 1999, Tsujino *et al* 2003, Rancati *et al* 2003, Jenkins *et al* 2003, Moiseenko *et al* 2003). It is likely that SOM_{dose} selected volumes above higher doses to complement the relatively lower MLD dose. It is also likely that SOM_{all}, in contrast to SOM_{dose}, did not select V_x features because they are more weakly complementary than the non-dose features. The distinction between MLD and V_x features is that MLD summarizes the entire DVH curve, whereas V_x is a single point on the curve. The frequently reported correlation of V_x to RP may be because of the association of V_x with the entire DVH shape. It is probable that the relative diversity of treatments in our database decreased V_x 's association with the DVH shape.

The non-dose features selected by model SOM_{all} are chemotherapy prior to RT (yes versus no), histology (squamous cell versus other histology) and tumor location (lower lobe versus other location). Our results indicate that patients who underwent chemotherapy prior to RT are more likely to have RP. Chemotherapy schedule has been previously reported as associated with the occurrence of pneumonitis (Theuws *et al* 1998a, 1998b, McDonald *et al* 1995). McDonald *et al* (1995) reported that some chemotherapeutic drugs can not only induce lung injury such as pneumonitis, but also enhance radiation-induced lung injury. Tumor location (lower lobe) was found to be associated with increased risk of pneumonitis. This has also been noted in the literature (Graham *et al* 1999, Hope *et al* 2006, Yorke *et al* 2002). Mice experiments have also suggested that tumor location is correlated to RP (Liao *et al* 1995, Travis *et al* 1997). In our model, patients with squamous cell lung cancer have higher risk of suffering RP (however, its correlation is weak; see discussion in the next subsection). To the best of our knowledge, there have been no reports from outside our institution associating tumor histology with RP (Das *et al* (2007), from our institution, also report this association in a decision tree model).

3.2. Evaluation of feature importance

Table 4 lists the input features selected by model SOM_{all}. The importance of each feature, measured by the AUC decrement resulting from its exclusion, is summarized in table 5. The most important features are approximately equivalent to the MLD (EUD $a = 0.9, 1.0$ and 1.1). Excluding these features resulted in the largest AUC drop of 0.12, from 0.73 to 0.61 ($p = 0.033$). The second most important factor in our model, chemotherapy prior to radiotherapy, resulted in an AUC decrement of 0.09 ($p = 0.041$). Exclusion of tumor location or histology resulted in only small drops in the AUC (AUC decrement $<0.01, p > 0.05$). Because neither tumor location nor histology is individually highly correlated with RP, their statistical significance is low. It is, however, not surprising that a more sophisticated model such as SOM selects input features with weak univariate correlation, since a factor that is not useful by itself can provide significant performance improvement when taken with others (Guyon and Elisseeff 2003).

3.3. Model testing and comparison

The cross-validated ROC curves are shown in figure 1 and figure 2 for models SOM_{all} and SOM_{dose}, respectively. The results of ROC analysis are shown in table 6. The AUC for training is notably lower than 1 because of the overfitting prevention resulting from the use of a training-evaluation set within the training set. The cross-validated AUC of model SOM_{all} (0.73) is higher than that of model SOM_{dose} (0.67), and the accuracy of model SOM_{all} (sensitivity: 71%, specificity: 68%) is better than that of model SOM_{dose} (sensitivity: 63%, specificity: 66%). The difference between the ROC areas from the two models was significant ($p=0.048$), indicating superiority of model SOM_{all} over model SOM_{dose}. Statistical significance also implies that the addition of non-dose factors improved the predictive ability of the SOM model.

The combination of multiple dose factors and non-dose factors into a single predictive model follows from previous works that have indicated synergism between dose and non-dose factors (Marks *et al* 1997, Lind *et al* 2002, Das *et al* 2007, Rodrigues *et al* 2004). Rodrigues *et al* (2004) reviewed 12 refereed publications and 2 conference abstracts with regard to three dose predictors: V_x , MLD and NTCP. Their review indicated that, although these three factors are associated with RP risk, they do not have high predictive power. This review suggested that a model based on several factors may improve our ability to predict the risk of clinically relevant RP. Marks *et al* (1997) state that ‘the use of NTCP/DVHs and patient-specific biologic factors (e.g. PFTs) represent the state-of-the-art in predictive measures of post-RT whole-lung function’, indicating possible synergy between dose and non-dose factors. Previous studies from our institution have suggested that combining a dose–volume factor with a pulmonary function test factor was more predictive than a model based on dose–volume factors alone (Marks *et al* 1997, Lind *et al* 2002). Lind *et al* (2006) observed that the area under the ROC curve for predictor V_{20} increased when the analysis was restricted to patients below 55 years of age.

3.4. Impact of data splitting on ten-fold cross-validation results

To evaluate the dependence of the cross-validated results on the assignment of patients to the ten groups, cross validation for SOM_{all} was run 200 times (each time, the data were randomly split ten-fold). The ROC curves for each of these 200 cross validations are shown superimposed in figure 3. The distribution of the areas under the 200 ROC curves is shown in figure 4 (mean = 0.724, standard deviation = 0.017). The small variance implies that the SOM_{all} model results are robust for ten-fold cross validation.

3.5. Biological mean lung dose (NTD_{mean})

The physical mean lung dose, MLD, does not take the effect of the dose per fraction into account. To gauge the impact of this effect in predicting RP, model SOM_{all} was run with ‘MLD’ (EUDs with exponents $a = 0.9, 1.0$ and 1.1) replaced by the biological mean lung dose, NTD_{mean}. The results with NTD_{mean} (cross-validated AUC: 0.72, sensitivity: 71%, specificity: 68%) were similar to those with ‘MLD’ (cross-validated AUC: 0.73, sensitivity: 68%, specificity: 71%). This would indicate that ‘MLD’ and NTD_{mean} may be considered as equivalent features in predicting RP.

3.6. Limitations on the applicability of the predictive model

The SOM model developed here to predict lung RP is only applicable within the scope imposed by the limitations of the database used to build the model (number and orientation of treatment fields, treatment field margins, total dose and fractionation). The database is limited to lung cancer patients who received radiation therapy via parallel opposed anterior-posterior fields followed by off-cord oblique fields. No gating technique was applied for

reduction of margins. Approximately 70% of patients were treated once daily at 1.8–2.0 Gy per fraction, and the remaining were treated twice daily at 1.6 Gy per fraction. Therefore, for example, the SOM model may not necessarily be applicable to lung cancer patients receiving intensity-modulated radiation therapy (IMRT) from a larger number of fields, where much higher lung volumes are usually irradiated to low dose and lower lung volumes receive high dose.

4. Conclusion

The self-organizing map (SOM) technique appears to be an effective, robust predictor of lung radiation pneumonitis. The model (SOM_{all}) developed from all factors is superior to the model (SOM_{dose}) developed from dose factors alone. The addition of non-dose factors improved model predictive ability. Among the selected input features in model SOM_{all}, the two most important features contributing to increased radiation pneumonitis risk are: doses approximately equivalent to the mean lung dose and chemotherapy prior to radiotherapy. Since the SOM model was built from the database at our institution, it may be poorly extrapolated to treatment techniques outside that contained in the database.

Acknowledgments

This work was supported by grants NIH R01 CA115748 and NIH R01 CA69579.

References

- Chang DT, Olivier KR, Morris CG, Liu C, Dempsey JF, Benda RK, Palta JR. The impact of heterogeneity correction on dosimetric parameters that predict for radiation pneumonitis. *Int. J. Radiat. Oncol. Biol. Phys* 2006;65:125–131. [PubMed: 16427214]
- Chen DR, Chang RF, Huang YL. Breast cancer diagnosis using self-organizing map for sonography. *Ultrasound Med. Biol* 2000;26:405–411. [PubMed: 10773370]
- Das SK, et al. Predicting radiotherapy-induced cardiac perfusion defects. *Med. Phys* 2005;32:19–27. [PubMed: 15719950]
- Das SK, Zhou S, Zhang J, Yin F, Dewhirst MW, Marks L. Predicting lung radiotherapy-induced pneumonitis using a model that combines parametric lyman probit with non-parametric decision trees. *Int. J. Radiat. Oncol. Biol. Phys* 2007;68:1212–1221. [PubMed: 17637394]
- Dubray B, Henrymar M, Meerwaldt JH, Noordijk EM, Dixon DO, Cosset JM, Thames HD. Radiation-induced lung damage after thoracic irradiation for Hodgkins disease—the role of fractionation. *Radiother. Oncol* 1995;36:211–217. [PubMed: 8532908]
- Graham MV, Purdy JA, Emami B, Harms W, Bosch W, Lockett MA, Perez CA. Clinical dose—volume histogram analysis for pneumonitis after 3D treatment for non-small cell lung cancer (NSCLC). *Int. J. Radiat. Oncol. Biol. Phys* 1999;45:323–329. [PubMed: 10487552]
- Green S, Weiss GR. Southwest-Oncology-Group Standard Response Criteria, End-Point Definitions and Toxicity Criteria. *Invest. New Drugs* 1992;10:239–253. [PubMed: 1487397]
- Guyon I, Elisseeff A. An introduction to variable and feature selection. *J. Mach. Learn. Res* 2003;3:1157–1182.
- Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer; 2002.
- Hernando ML, et al. Radiation-induced pulmonary toxicity: a dose–volume histogram analysis in 201 patients with lung cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2001;51:650–659. [PubMed: 11597805]
- Hope AJ, Lindsay PE, ElNaqa I, Alaly JR, Vicic M, Bradley JD, Deasy JO. Modeling radiation pneumonitis risk with clinical, dosimetric, and spatial parameters. *Int. J. Radiat. Oncol. Biol. Phys* 2006;65:112–124. [PubMed: 16618575]
- Jenkins P, D'Amico K, Benstead K, Elyan S. Radiation pneumonitis following treatment of non-small cell lung cancer with continuous hyperfractionated accelerated radiotherapy (CHART). *Int. J. Radiat. Oncol. Biol. Phys* 2003;56:360–366. [PubMed: 12738310]

- Kocak Z, Evans ES, Zhou SM, Miller KL, Folz RJ, Shafman TD, Marks LB. Challenges in defining radiation pneumonitis in patients with lung cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2005a;62:635–638. [PubMed: 15936538]
- Kocak Z, Yu X, Zhou SM, D'Amico TA, Hollis D, Kahn D, Tisch A, Shafman TD, Marks LB. The impact of pre-radiotherapy surgery on radiation-induced lung injury. *Clin. Oncol* 2005b;17:210–216.
- Kohonen, T. *Self-Organizing Maps*. New York: Springer; 1995.
- Kong F-M, et al. Final toxicity results of a radiation-dose escalation study in patients with non-small cell lung cancer (nscLc): predictors for radiation pneumonitis and fibrosis. *Int. J. Radiat. Oncol. Biol. Phys* 2006;65:1075–1086. [PubMed: 16647222]
- Kwa SLS, et al. Radiation pneumonitis as a function of mean lung dose: an analysis of pooled data of 540 patients. *Int. J. Radiat. Oncol. Biol. Phys* 1998;42:1–9. [PubMed: 9747813]
- Lebesque JV, Keus RB. The simultaneous boost technique—the concept of relative normalized total dose. *Radiother. Oncol* 1991;22:45–55. [PubMed: 1947212]
- Liao ZX, Travis EL, Tucker SL. Damage and morbidity from pneumonitis after irradiation of partial volumes of mouse lung. *Int. J. Radiat. Oncol. Biol. Phys* 1995;32:1359–1370. [PubMed: 7635776]
- Lind PA, Marks LB, Hollis D, Fan M, Zhou SM, Munley MT, Shafman TD, Jaszczak RJ, Coleman RE. Receiver operating characteristic curves to assess predictors of radiation-induced symptomatic lung injury. *Int. J. Radiat. Oncol. Biol. Phys* 2002;54:340–347. [PubMed: 12243806]
- Lind PA, Wennberg B, Gagliardi G, Rosfors S, Blom-Goldman U, Lidestahl A, Svane G. ROC curves and evaluation of radiation-induced pulmonary toxicity in breast cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2006;64:765–770. [PubMed: 16257129]
- Luijk PV, Faber H, Meertens H, Kampinga H, Schippers JM, Coppes R. Sub-clinical heart damage enhances radiation-induced lung function loss. *Int. J. Radiat. Oncol. Biol. Phys* 2005;63:S460–S.
- Lyman JT. Complication probability as assessed from dose volume histograms. *Radiat. Res* 1985;104:S13–S19.
- Markey MK, Lo JY, Tourassi GD, Floyd CE. Self-organizing map for cluster analysis of a breast cancer database. *Artif. Intell. Med* 2003;27:113–127. [PubMed: 12636975]
- Marks LB. Dosimetric predictors of radiation-induced lung injury. *Int. J. Radiat. Oncol. Biol. Phys* 2002;54:313–316. [PubMed: 12243802]
- Marks LB, et al. Physical and biological predictors of changes in whole-lung function following thoracic irradiation. *Int. J. Radiat. Oncol. Biol. Phys* 1997;39:563–570. [PubMed: 9336133]
- Martel MK, Tenhaken RK, Hazuka MB, Turrisi AT, Fraass BA, Lichter AS. Dose–volume histogram and 3D treatment planning evaluation of patients with pneumonitis. *Int. J. Radiat. Oncol. Biol. Phys* 1994;28:575–581. [PubMed: 8113100]
- McDonald S, Rubin P, Phillips TL, Marks LB. Injury to the lung from cancer therapy—clinical syndromes, measurable end points, and potential scoring systems. *Int. J. Radiat. Oncol. Biol. Phys* 1995;31:1187–1203. [PubMed: 7713782]
- Metz CE, Herman BA, Roe CA. Statistical comparison of two ROC-curve estimates obtained from partially-paired datasets. *Med. Decis. Making* 1998;18:110–121. [PubMed: 9456215]
- Moiseenko V, Craig T, Bezjak A, Van Dyk J. Dose–volume analysis of lung complications in the radiation treatment of malignant thymoma: a retrospective review. *Radiother. Oncol* 2003;67:265–274. [PubMed: 12865174]
- Niemierko A. A generalized concept of equivalent uniform dose (EUD). *Med. Phys* 1999;26:1100.
- Rancati T, Ceresoli GL, Gagliardi G, Schipani S, Cattaneo GM. Factors predicting radiation pneumonitis in lung cancer patients: a retrospective study. *Radiother. Oncol* 2003;67:275–283. [PubMed: 12865175]
- Robnett TJ, Machtay M, Vines EF, McKenna MG, Algazy KM, McKenna WG. Factors predicting severe radiation pneumonitis in patients receiving definitive chemoradiation for lung cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2000;48:89–94. [PubMed: 10924976]
- Rodrigues G, Lock M, D'Souza D, Yu E, Van Dyk J. Prediction of radiation pneumonitis by dose–volume histogram parameters in lung cancer—a systematic review. *Radiother. Oncol* 2004;71:127–138. [PubMed: 15110445]

- Schallenkamp JM, Miller RC, Brinkmann DH, Foote T, Garces YI. Incidence of radiation pneumonitis after thoracic irradiation: dose–volume correlates. *Int. J. Radiat. Oncol. Biol. Phys* 2007;67:410–416. [PubMed: 17236964]
- Seppenwoolde Y, Lebesque JV, de Jaeger K, Belderbos JSA, Boersma LJ, Schilstra C, Henning GT, Hayman JA, Martel MK, Ten Haken RK. Comparing different NTCP models that predict the incidence of radiation pneumonitis. *Int. J. Radiat. Oncol. Biol. Phys* 2003;55:724–735. [PubMed: 12573760]
- Setiono R. Feed forward neural network construction using cross validation. *Neural Comput* 2001;13:2865–2877. [PubMed: 11705414]
- Swets, J.; Pickett, R. *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*. New York: Academic; 1982.
- Theuvs JCM, Kwa SLS, Wagenaar AC, Boersma LJ, Damen EMF, Muller SH, Baas P, Lebesque JV. Dose-effect relations for early local pulmonary injury after irradiation for malignant lymphoma and breast cancer. *Radiother. Oncol* 1998a;48:33–43. [PubMed: 9756170]
- Theuvs JCM, Kwa SLS, Wagenaar AC, Seppenwoolde Y, Boersma LJ, Damen EMF, Muller SH, Baas P, Lebesque JV. Prediction of overall pulmonary function loss in relation to the 3D dose distribution for patients with breast cancer and malignant lymphoma. *Radiother. Oncol* 1998b; 49:233–243. [PubMed: 10075256]
- Travis EL, Liao ZX, Tucker SL. Spatial heterogeneity of the volume effect for radiation pneumonitis in mouse lung. *Int. J. Radiat. Oncol. Biol. Phys* 1997;38:1045–1054. [PubMed: 9276371]
- Tsujino K, Hirota S, Endo M, Obayashi K, Kotani Y, Satouchi M, Kado T, Takada Y. Predictive value of dose–volume histogram parameters for predicting radiation pneumonitis after concurrent chemoradiation for lung cancer. *Int. J. Radiat. Oncol. Biol. Phys* 2003;55:110–115. [PubMed: 12504042]
- Tsujino K, Hirota S, Kotani Y, Kado T, Yoden E, Fujii O, Soejima T, Adachi S, Takada Y. Radiation pneumonitis following concurrent accelerated hyperfractionated radiotherapy and chemotherapy for limited-stage small-cell lung cancer: dose–volume histogram analysis and comparison with conventional chemoradiation. *Int. J. Radiat. Oncol. Biol. Phys* 2006;64:1100–1105. [PubMed: 16373082]
- Yorke ED, Jackson A, Rosenzweig KE, Merrick SA, Gabrys D, Venkatraman ES, Burman CM, Leibel SA, Ling CC. Dose–volume factors contributing to the incidence of radiation pneumonitis in non-small cell lung cancer patients treated with three-dimensional conformal radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys* 2002;54:329–339. [PubMed: 12243805]

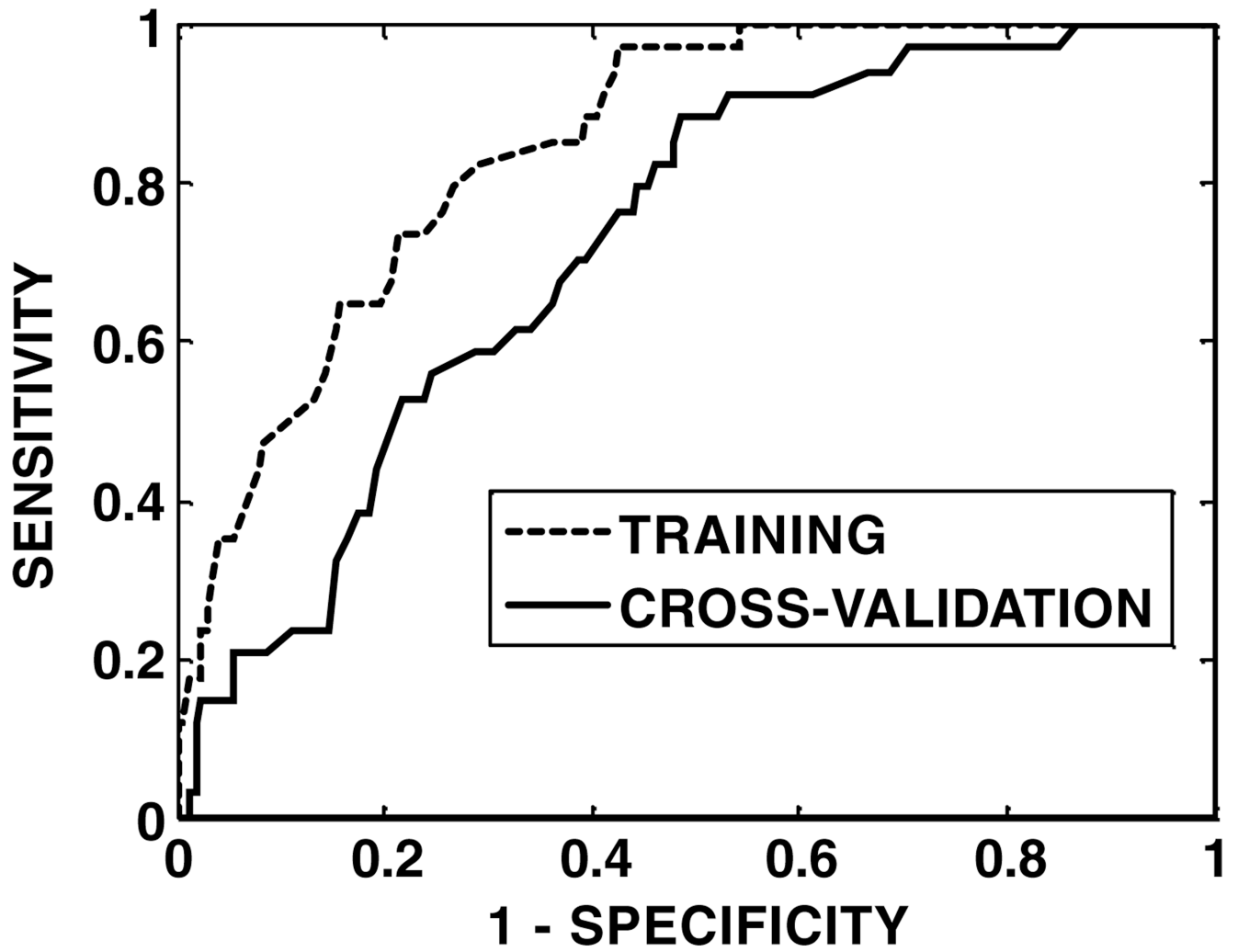


Figure 1. ROC curves for the model SVM_{all} on training (ROC area = 0.85) and cross-validation (ROC area = 0.73) data.

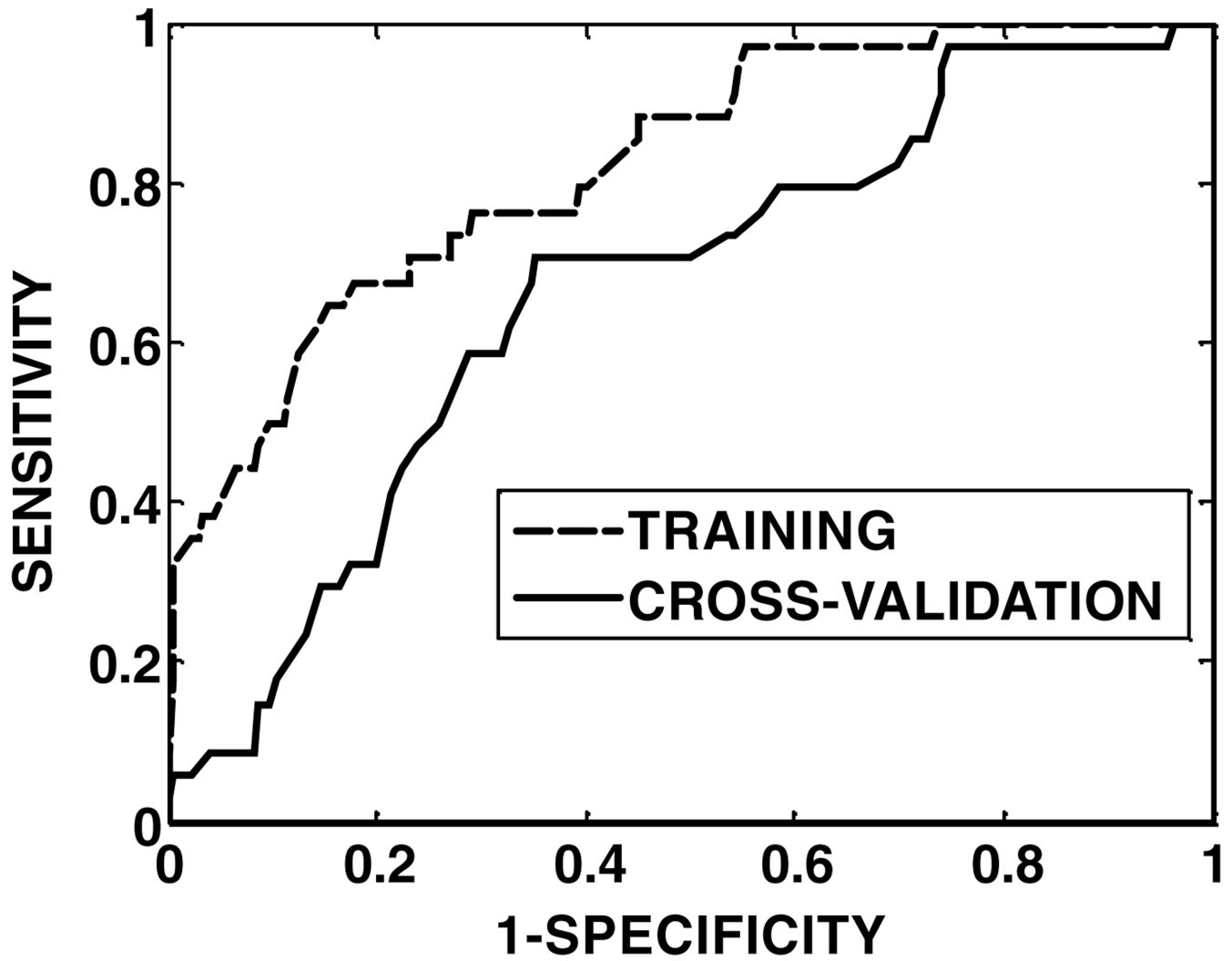


Figure 2.
ROC curves for the model SVM_{dose} on training (ROC area = 0.82) and cross-validation (ROC area = 0.67) data.

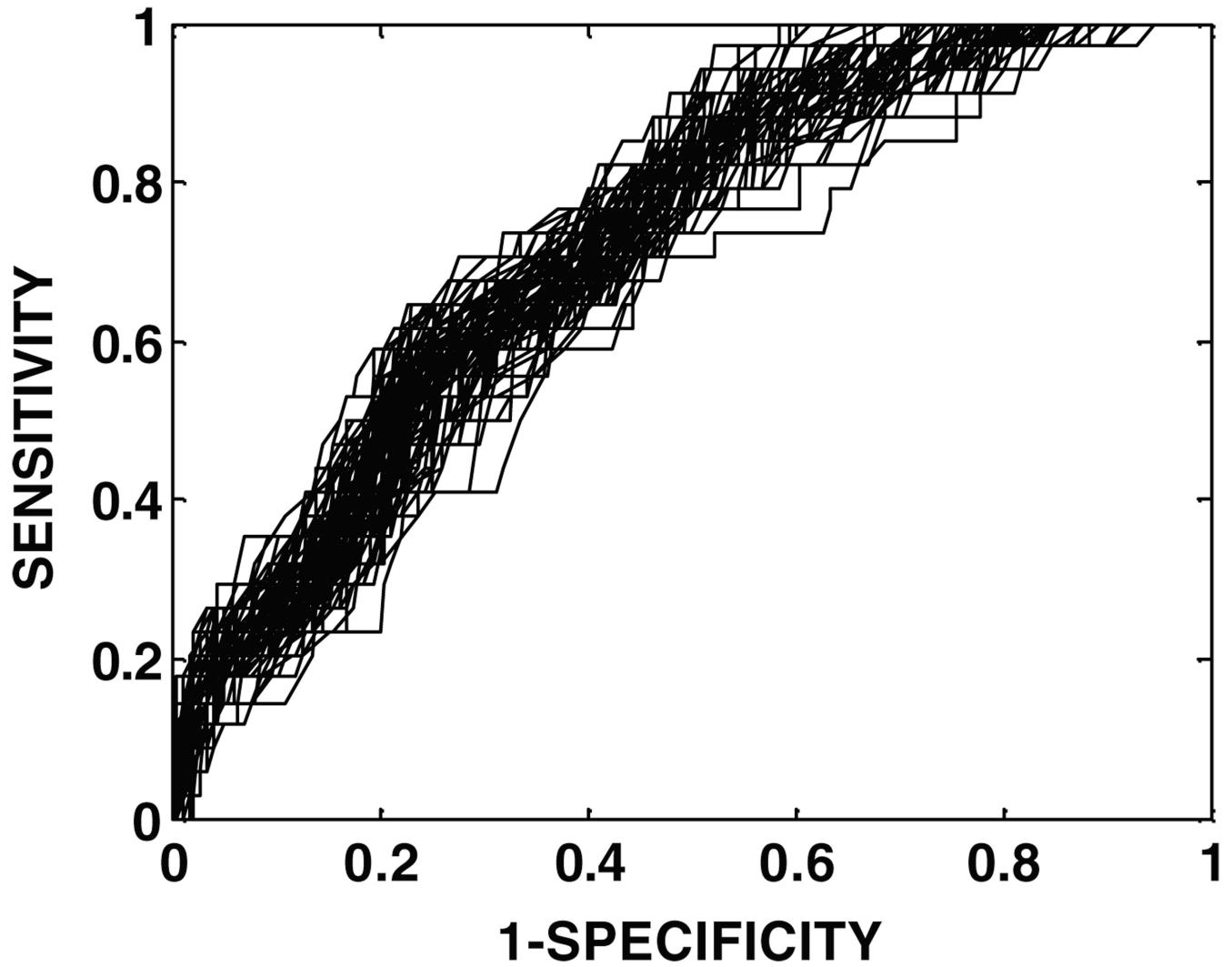


Figure 3.
The cross-validated ROC curves for model SVM_{all} , for 200 random splits of the ten cross-validation groups (area under ROC curves = 0.724 ± 0.017).

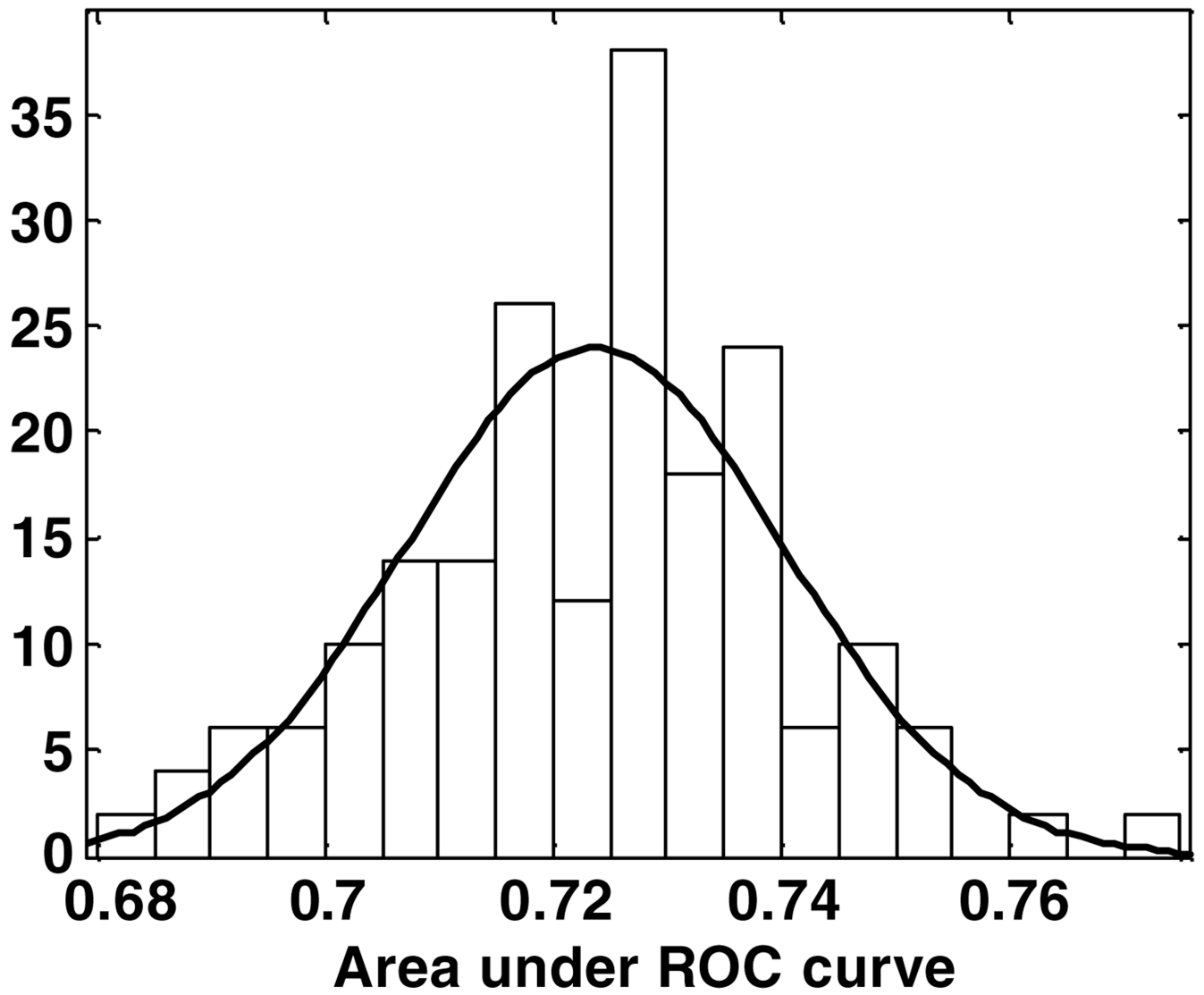


Figure 4. The distribution of the areas under the cross-validated ROC curves, for model SVM_{all} , for 200 random splits of the ten cross-validation groups (mean = 0.724, standard deviation = 0.017).

Table 1

Patient and treatment characteristics.

Characteristic	Percentage of all patients or mean (range)
Radiation pneumonitis	15.5
No radiation pneumonitis	84.5
Age (years)	64 (27–87)
Gender: male/female	54/56
Race: white/others	83/17
Pre-RT FEV1%	64 (15–127)
Pre-RT DLCO%	63 (11–129)
Fractionation: once/twice daily	70/30
RT dose (Gy)	
Once daily treatment	63.0 (30.0–86.4)
Twice daily treatment	72.3 (45.0–86.4)
Histological type	
SCLC	11
NSCLC (squamous/adeno/other)	88 (27/19/42)
Other	1
Tumor stage	
I–II	17
III–IV	81
Recurrent	2
Tumor location	
Central/peripheral	61/39
Upper/lower/middle	60/22/18
Right/left location	58/42
Chemotherapy schedule (relative to RT)	
Concurrent with RT	14
Pre-RT	48
Pre-RT + concurrent RT	6
Concurrent RT+ post-RT	3
Post-RT	1
No chemotherapy	28

SCLC, small cell lung cancer; NSCLC, non-small cell lung cancer; FEV1, forced expiratory volume in 1 second; DLCO, diffusion capacity of carbon monoxide.

Table 2

Radiation pneumonitis grading system.

Grading system	Grade	Definition
The modified National Cancer Institute Common Toxicity Criteria (NCICTC)	0	No increase in symptoms
	1	Symptoms not requiring initiation or increase in steroids and/or oxygen
	2	Symptoms requiring initiation or increase in steroids
	3	Symptoms requiring oxygen
	4	Symptoms requiring assisted ventilation or causing death.
Southwest Oncology Group Toxicity Criteria	1	Radiographic changes are observed but the symptoms do not require steroids
	2	Requiring steroids
	3	Oxygen required
	4	Assisted ventilation is necessary
	5	Fatal

Table 3

The area under the ROC curve (AUC) of models SOM_{all} and SOM_{dose} , for various SOM topologies.

Topology of SOM	AUC of SOM_{all}		AUC of SOM_{dose}	
	Training construction	Training evaluation	Training construction	Training evaluation
3×2	0.76	0.67	0.72	0.63
3×3	0.81	0.73	0.78	0.66
4×3	0.85	0.82	0.82	0.74
4×4	0.87	0.78	0.85	0.72

Table 4

The input features selected by model SOM_{all} and SOM_{dose} . V_x is the lung volume above x Gy. Higher values of the dose variables are indicative of greater risk of radiation pneumonitis. Arrows next to the non-dose variables indicate higher (\uparrow) or lower (\downarrow) risk.

SOM_{all}	SOM_{dose}
EUD $a = 0.9$	EUD $a = 0.7$
EUD $a = 1.0$	EUD $a = 0.8$
EUD $a = 1.1$	EUD $a = 0.9$
Chemotherapy prior to RT (yes \uparrow versus no \downarrow)	EUD $a = 1.0$
Histology (squamous cell \uparrow versus other histology \downarrow)	V_{40}
Tumor location (lower lobe \uparrow versus other location \downarrow)	V_{42}
	V_{44}

Table 5

Evaluation of importance of input features by individual exclusion.

Input feature excluded	AUC decrement	p-Value
EUD $\alpha = 0.9, 1.0, 1.1$	0.12	0.033
Chemotherapy prior to radiotherapy (yes or no)	0.09	0.041
Tumor position (lower lobe or not)	<0.01	0.166
Histology (squamous cell or not)	<0.01	0.236

Table 6The results of ROC analysis for model SOM_{all} and SOM_{dose} .

		SOM_{all}	SOM_{dose}
Training	AUC	0.85	0.82
	Sensitivity	76%	71%
	Specificity	76%	74%
Cross validation	AUC	0.73	0.67
	Sensitivity	71%	63%
	Specificity	68%	66%