Acta Oncologica, 2015; Early Online: 1–8

informa
healthcare

ORIGINAL ARTICLE

Predicting radiation-induced valvular heart damage

LAURA CELLA[1,2], JUNG HUN OH[3], JOSEPH O. DEASY[3], GIUSEPPE PALMA[1], RAFFAELE LIUZZI[1,2], VITTORIA D'AVINO[1], MANUEL CONSON[1,2], MARCO PICARDI[4], MARCO SALVATORE[2] & ROBERTO PACELLI[1,2]

[1]*Institute of Biostructure and Bioimaging, National Council of Research (CNR), Naples, Italy,* [2]*Department of Advanced Biomedical Sciences, Federico II University School of Medicine, Naples, Italy,* [3]*Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York, USA and* [4]*Department of Clinical Medicine and Surgery, Federico II University School of Medicine, Naples, Italy*

ABSTRACT

Purpose. To develop a predictive multivariate normal tissue complication probability (NTCP) model for radiation-induced heart valvular damage (RVD). The influence of combined heart-lung irradiation on RVD development was included.

Material and methods. Multivariate logistic regression modeling with the least absolute shrinkage and selection operator (LASSO) was used to build an NTCP model to predict RVD based on a cohort of 90 Hodgkin lymphoma patients treated with sequential chemo-radiation therapy. In addition to heart irradiation factors, clinical variables, along with left and right lung dose-volume histogram statistics, were included in the analysis. To avoid overfitting, 10-fold cross-validation (CV) was used for LASSO logistic regression modeling, with 50 reshuffled cycles. Model performance was assessed using the area under the receiver operating characteristic (ROC) curve (AUC) and Spearman's correlation coefficient (Rs).

Results. At a median follow-up time of 55 months (range 12–92 months) after the end of radiation treatment, 27 of 90 patients (30%) manifested at least one kind of RVD (mild or moderate), with a higher incidence of left-sided valve defects (64%). Fourteen prognostic factors were frequently selected (more than 100/500 model fits) by LASSO, which included mainly heart and left lung dosimetric variables along with their volume variables. The averaged cross-validated performance was AUC-CV = 0.685 and Rs = 0.293. The overall performance of a final NTCP model for RVD obtained applying LASSO logistic regression to the full dataset was satisfactory (AUC = 0.84, Rs = 0.55, $p < 0.001$).

Conclusion. LASSO proved to be an improved and flexible modeling method for variable selection. Applying LASSO, we showed, for the first time, the importance of jointly considering left lung irradiation and left lung volume size in the prediction of subclinical radiation-related heart disease resulting in RVD.

Late cardiovascular effects of thoracic irradiation are the result of coronary, myocardial and valvular pathology [1]. Radiation-induced heart valvular defects (RVD) may be regarded as a first step from an asymptomatic stage to overt heart failure, and as such may be considered as an early indicator of some types of late heart damage [2]. A better understanding of the underlying mechanism of such toxicity could help in developing protective strategies to avoid clinically relevant cardiovascular side effects. Experimental studies in rats showed that irradiation of heart, lungs, or both independently induces specific cardiac dysfunction and associated pulmonary vascular damage, in a negative synergy [3,4]. However, clinical studies in radiotherapy patients are necessary in order to link these results to humans [5].

In our latest analysis on Hodgkin lymphoma (HL) survivors [6–8] investigating RVD, we found that the risk of RVD cannot be modeled using traditional normal tissue complication probability (NTCP) models only based on heart dose-volume distribution. A multivariate logistic regression approach showed superior predictive power compared to traditional NTCP models and an improved

Correspondence: L. Cella, Institute of Biostructures and Bioimaging, National Council of Research (CNR), 80131 Naples, Italy. Tel: +39 081 2203427. E-mail: laura.cella@cnr.it

(Received 24 October 2015; accepted 3 February 2015)

ISSN 0284-186X print/ISSN 1651-226X online © 2015 Informa Healthcare
DOI: 10.3109/0284186X.2015.1016624

performance can be obtained with the inclusion of heart and lung volume terms, indicating that heart-lung interactions are apparently important for this endpoint. For radiation-induced lung injury, in contrast, we found that left lung irradiation, compared to right lung irradiation, had a greater impact on lung toxicity. To better clarify the influence of combined heart-lung irradiation on RVD, in the present analysis we extend our investigation including more clinical and dosimetric factors. We hypothesize that incorporating left and right lung dose-volume histogram (DVH) factors separately in addition to heart DVH factors would result in a better combination of predictor variables. Further improvement of the current work is to use the least absolute shrinkage and selection operator (LASSO) method with cross-validation (CV) parameter tuning.

Recently, improved statistical machine learning methods such as LASSO, have been applied to build multivariate NTCP models [9–11]. In their comparative study, Xu et al. [10] showed that LASSO models have significantly better predictive power than the stepwise selection method in predicting Grade 2 or higher xerostomia. In another study [11] using the same dataset, Xu et al. confirmed the usefulness of LASSO in conjunction with repeated double CV and permutation tests for the validation of NTCP models. Lee et al. [9] analyzed quality of life questionnaire data for head and neck cancer (HNC) patients to build a multivariate LASSO-based NTCP model for the risk of xerostomia after radiotherapy, which resulted in satisfactory performance. One of the main advantages of LASSO is that it allows us to perform feature selection and model construction simultaneously. The success of LASSO in NTCP modeling and its usefulness inspired us to develop LASSO-based NTCP models for the risk of RVD.

Material and methods

Patient dataset

A total of 115 consecutive HL survivors who were treated with post-chemotherapy (post-CHT) supradiaphragmatic involved-field radiation therapy (RT) at our institution from November 2001 until November 2012 were identified. The study selection criteria included: 1) cardiac evaluation by electrocardiography and echocardiography before CHT, after CHT and before RT, and periodically after RT; 2) no pre-treatment cardiac heart disease; 3) a follow-up of at least 36 months; and 4) availability of three-dimensional (3D) dose maps.

Twelve of 115 (10.4%) patients had some type of pre-treatment cardiac abnormalities and were excluded from further evaluation. Five patients were excluded because 3D dose maps were not retrievable, and eight patients were excluded for having a follow-up <36 months (supplementary Figure 1, available online at: http://informahealthcare.com/doi/abs/ 10.3109/0284186X.2015.1016624). After exclusion, a total of 90 patients were included with a median follow-up of 80 months (range 38–140 months). Complete description of clinical variables and characteristics of the patients eligible for the analysis is reported in Table 1 and supplementary Table 1, available online at: http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1016624. At a median time of 55 months (range 12–92 months) after the end of radiation treatment, 27 (30.0%) of 90 patients manifested at least one kind of RVD (mild or moderate). Higher incidence of left-sided valve defects was confirmed: 20 patients (22.2%) had defects at mitral valve, five patients (5.6%) at aortic valve, and 14 patients (15.6%) at tricuspid valve. No pulmonary valve toxicity was recorded. Some patients had more than one valve defect.

All the characteristics of the study design have been described in detail in our previous publication [8]. The present study represents an update on an extended dataset including 90 patients compared to previous 56 patients. Briefly, a diagnosis of RVD was based on the presence of regurgitation and/or stenosis (mild, moderate, or severe) of at least one of the aortic, mitral, tricuspid and pulmonary valves. All patients were treated with full 3D free-breathing computed tomography (CT)-based radiation treatment planning. RT was administered using 6–20 MV photon beams from a linear accelerator with antero-posterior-posteroanterior fields. A median total dose of 32 Gy (range 21–41 Gy) was prescribed. The daily fraction size ranged between 1.5 Gy and 1.8 Gy.

For all patients, contoured structures included the whole heart, and the left and right lung tissues, retrospectively contoured on planning CTs by the same radiation oncologist, following the heart atlas contouring guidelines published by Feng et al. [12], as well as the RTOG 1106 recommendations [13]. The total lung, combined with the left and right lung, was also considered.

Dose volume histogram metrics

DVH data were extracted from treatment planning data using the open-source software CERR (Computational Environment for Radiotherapy Research) [14]. Individual DICOM-RT plans (doses and organ contours) were converted to Matlab-readable format for further analysis (MATLAB version 7.6. Mathworks, Natik, MA, USA). Various DVH metrics were extracted for modeling: the minimum dose to x% highest dose volume (D_x); the percentage volume

Table I. Clinical variables and univariate analysis for radiation-induced heart valvular damage RVD incidence.

Characteristic			Univariate analysis	
			Rs	p-Value
Continuous variables	Median (range)			
Age (year)	28 (14–70)		0.011	0.924
Body mass index (kg/m^2)	25.0 (15.1–39.3)		0.174	0.116
Heart volume (cm^3)	539.3 (336.0–898.4)		0.231	0.028
Lungs volume (cm^3)	2708.4 (1396.1–5729.4)		−0.224	0.034
Left lung volume (cm^3)	1254.9 (630.6–2792.5)		−0.280	0.008
Right lung volume (cm^3)	1465.2 (732.0–2989.5)		−0.162	0.126
Categorical variables	N	%		
Gender				
Female	51	56.7		
Male	39	43.3	−0.132	0.215
Histology				
Nodular sclerosis	64	71.1		
Mixed cellularity	21	23.3		
Lymphocyte-rich-classical	5	5.5	0.063	0.555
Stage				
I–II	70	77.8		
III–IV	20	22.2	−0.058	0.585
Chemotherapy regimen				
ABVD	20	22.2		
VEBEP	70	77.8	−0.058	0.585
Risk factors*				
None	64	71.1		
Yes	26	28.9	0.096	0.367
Smokers	12			
Diabetes	2			
Hypertension	5			
Hypercholesterolemia	1			
Lung disease	6			
Thyroid disease	9			

ABVD, doxorubicin, bleomycin, vinblastine, and dacarbazine; CRP, C-reactive protein; LDH, lactate dehydrogenase; VEBEP, vinblastine, etoposide, bleomycin, epidoxorubicin, and prednisone.
*There can be multiple factors simultaneously.

receiving at least x dose (V_x); the absolute volume receiving at least x dose (AV_x); maximum dose (D_{max}); and mean dose (D_{mean}).

Complete DVHs and clinical data analyzed in the present study are available at http://www.ibb.cnr.it/?command = viewcms&id = 88.

Predictive model building

First, univariate logistic regression analysis for each candidate prognostic (clinical and dosimetric) factor was performed based on Spearman's rank correlation coefficient (Rs). Only the variables most highly correlated with RVD (with a cut-off of $Rs > 0.21$) were included in the subsequent analysis. In order to avoid a collinearity problem, among these variables, redundant variables were further removed such that, among variables highly correlated with each other (Pearson's correlation coefficient larger than 0.85), the variables with the lower correlation with RVD were removed.

We previously developed multivariate logistic regression methods for NTCP models with bootstrapping or leave-one-out cross-validation (LOOCV) approaches in a manner of forward feature selection [15]. In this study, for RVD modeling, we employed logistic regression and LASSO logistic regression methods. The LASSO is a linear regression model that employs penalized estimation to favor sparse models with a small number of variables. A key advantage of LASSO model is that it can simultaneously perform feature selection and model construction. LASSO logistic regression is a LASSO method for binary outcomes as in the case of our study.

For a given dataset of n patients and m variables, the logistic regression model is defined as:

$$NTCP_i\left(RVD|x_i\right) = \frac{1}{1+e^{-y_i}} \tag{1}$$

with

$$y_i = \beta_0 + \sum_{k=1}^{m} \beta_k x_{ik} \tag{2}$$

Where β_k are the regression coefficients corresponding to x_{ik} prognostic factors for the *i-th* patient.

Let $p(x_i) = NTCP(RVD\ x_i)$, and r_i is either 1 or 0 depending on whether a patient had RVD. Adding the L1 regularization penalty term, we have the following maximum log-likelihood equation:

$$\max_{(\beta_0, \beta) \in R^{p+1}} \left[\sum_{i=1}^{n} r_i \log p(x_i) + (1 - r_i) \log (1 - p(x_i)) - \lambda \sum_{k=1}^{m} |\beta_k| \right] \quad (3)$$

where λ is a non-negative tuning parameter that controls the sparsity of the model, determined by CV.

To avoid overfitting, a 10-fold CV approach was used for both logistic regression and LASSO logistic regression modeling. In the approach, 90 patients were randomly shuffled and were split into 10 groups (folds). In the first iteration, the first group was used for testing while the remaining nine groups for training. In the second iteration, the second group was used for testing and the remaining nine groups for training, and so on until the last group was used for testing. This entire procedure was iterated 50 times with data shuffling, resulting in 500 different models. For LASSO logistic regression modeling, at each iteration, a nested 10-fold CV (splitting 9 training groups into new 10 groups) was used to find the optimal λ value (λ_0), i.e. the λ value that minimizes the deviance curve. As the λ value increases, more coefficients are set to zero, leading to sparser model. Prognostic factors selected by λ_0 or applying the "one standard error rule" (that chooses λ_1 as the largest λ value within one standard error of the model corresponding to λ_0) were used for building NTCP models [11]. Standardization of variables was performed before solving the LASSO problem in order to guarantee the penalty term would treat variables in a comparable way.

Model performance was evaluated by the area under the receiver operating characteristic (ROC) curve (AUC) and Rs. Calibration plots were generated for graphical assessment of the agreement between observed outcomes and predictions.

Results

Applying the variable exclusion criteria described in the statistical modeling section, we were left with 26 candidate prognostic factors. The candidate variables included 22 dosimetric variables and four clinical variables (i.e. heart, lungs, left and right lung volumes). The univariate analysis for representative D_x and V_x parameters and Rs plots of RVD incidence as a function of D_x and V_x for heart, lungs, left and right lung are reported in Supplementary Table II and Supplementary Figure 2, available online at: http://informahealthcare.com/doi/abs/ 10.3109/0284186X.2015.1016624.

Fourteen prognostic factors were selected by LASSO more than 100 times of 500 sub-iterations (20%). Making the threshold as low as 30/500 does not change factors selected, indicating that factor selection was relatively robust. Variables selected by LASSO, besides heart dosimetric parameters and heart volume, included mainly left lung dosimetric parameters along with left lung volume (Figure 1a and b).

The LASSO logistic regression model resulted in an averaged AUC of 0.69 [standard deviation (SD): 0.022] and an averaged Rs of 0.29 (SD: 0.035) [averaged $p = 0.008$ (SD: 0.008)], whereas the logistic regression model resulted in an averaged AUC of 0.68 (SD: 0.03) and an averaged Rs of 0.23 (SD: 0.04) [averaged $p = 0.013$ (SD: 0.016)]. The performance of the LASSO logistic regression model was slightly better than that of the logistic regression model (Mann-Whitney U-test, $p = 0.04$). The LASSO method also had a smaller variability than logistic regression (Figure 2).

The 14 frequently selected prognostic factors were used to find a final NTCP model for RVD applying LASSO logistic regression. A model with 12 variables was found at λ_0: four heart-related variables (D10, Dmax, AV30 and volume), five left lung-related variables (D15, D25, D95, Dmax, volume), two right lung-related variables (D25 and Dmax) and lungs Dmax. Of these, a subset of eight variables was selected at λ_1 (Figure 1c and d). Selected variables and the corresponding coefficients along with the models' performance measures were reported in Table II. To generalize the application of the LASSO models, we also reported the model coefficients for untransformed variables.

By analyzing both the entire cohort of patients and only the group of patients who developed RVD, no difference was found between left lung and right lung DVHs (Supplementary Figures 3a-d, available online at: http://informahealthcare.com/doi/abs/ 10.3109/0284186X.2015.1016624), meaning that both lungs were equally irradiated.

To investigate the cross-variable interaction between heart-related variables and lung-related variables, we designed a simple two-variable logistic regression model adding an interaction term to the regression model. We hypothesized that if there is correlation between heart and lung, a logistic regression model with two variables would show a higher predictive performance using a lung variable, a heart variable and a term in which the former variables are multiplied. Let A and B denote two variables. In such a model, three variables A, B and A$\times$B were input, and AUC and Rs were

(a)

No	Variable	No	Variable	No	Variable
1	D5 left lung	11	D95 lungs	21	**AV30 heart**
2	**D15 left lung**	12	V30 lungs	22	**Dmax heart**
3	**D25 left lung**	13	V35 lungs	23	**Heart volume**
4	**D95 left lung**	14	AV35 lungs	24	**Left lung volume**
5	**Dmax left lung**	15	**Dmax lungs**	25	Right lung volume
6	**D25 right lung**	16	**D10 heart**	26	Lungs volume
7	V30 right lung	17	D35 heart		
8	**Dmax right lung**	18	**V30 heart**		
9	D5 lungs	19	AV20 heart		
10	D20 lungs	20	**AV25 heart**		

Abbreviations: AVX = absolute volume receiving x Gy, Dmax = maximum dose, Dx = the minimum dose to x% highest dose volume, Vx = percentage volume receiving x Gy, Frequently (>100) selected variables are in bold.

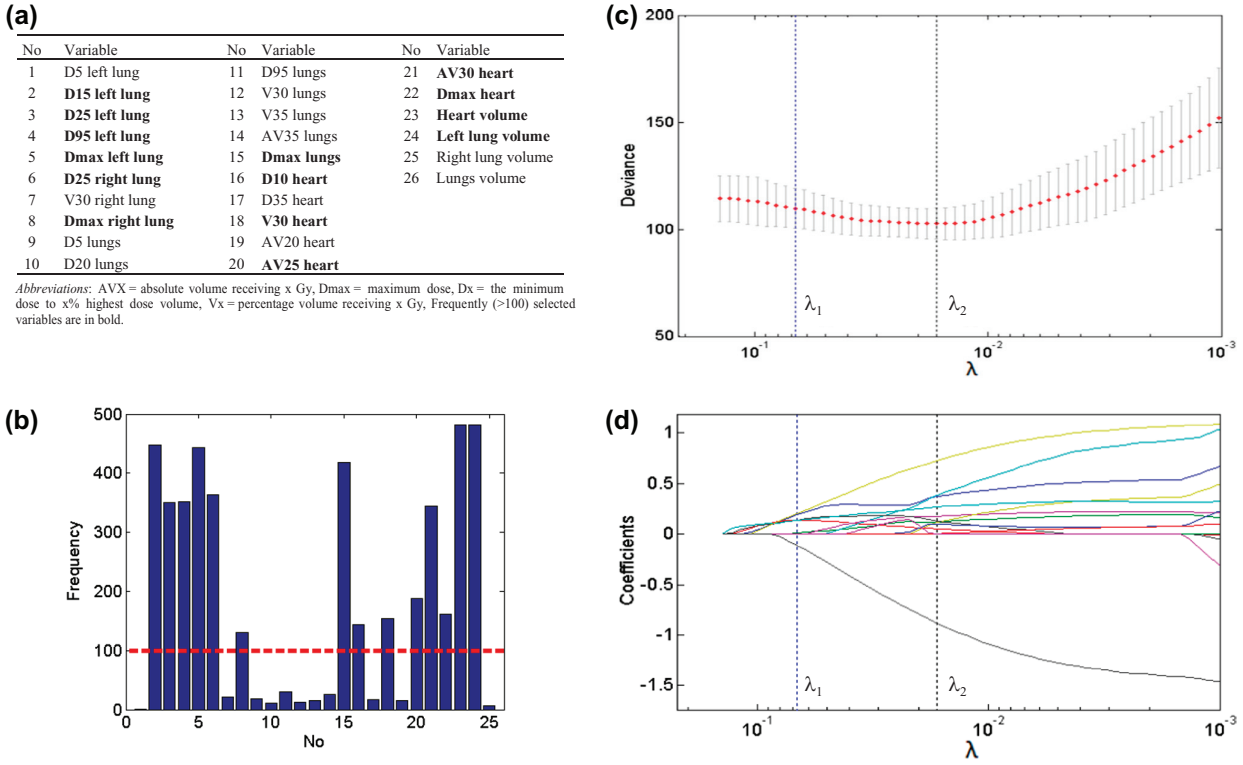

Figure 1. (a) List of the 26 candidate variables and (b) their selection frequencies by least absolute shrinkage and selection operator (LASSO) logistic regression models during 10-fold cross validation with 50 iterations. Variables are named by their index numbers as listed in (a). The red dashed-line indicates the frequency of occurrence of 100; 14 variables are selected more than 100 times. (c) Deviance values of the LASSO models as a function of the regularization parameter λ. The optimal penalty λ, determined by 10-fold cross-validation, is the value that minimizes the deviance curve. The plot identifies the minimum-deviance (dashed line λ_0) and the minimum deviance within one standard error (dashed line λ_1) from λ_0. (d) Trace plot showing non-zero model coefficients as a function of the regularization parameter λ. As λ increases to the left, LASSO sets various coefficients to zero, removing them from the model. When λ corresponds to the minimum-deviance (dashed line λ_0), 12 variables are selected. When λ corresponds to the minimum-deviance within one standard error (dashed line λ_1) from λ_0, 8 variables are selected.

calculated (Supplementary Table III available online at: http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1016624). As expected, models with a combination of heart and left lung variables showed the better performance as confirmed by higher AUC values.

Discussion

Trends towards longer survival following cancer therapy mean that more importance has to be given to iatrogenic effects such as radiation-induced toxicity [16]. Individualized patient risk assessment should be achieved in order to guide treatment choice and appropriate follow-up. The aim of the present work was to investigate all possible patient-related clinical and dosimetric factors that could contribute to the development of radiation-induced valvular defects within a cohort of 90 HL survivors. This endpoint has two peculiarities: RVD may be considered as an early indicator of late heart damage [2] and RVD cannot be explained by vascular damage because the valves lack vasculature. Possibly, this damage is consequential to late injury of the surrounding myocardial endothelium, leading to fibrosis [17]. In addition, many studies have reported on the higher incidence of defects to left-sided valves [18–20]. Left-sided valvular defects are more common also in the natural population; it has been argued that a higher pressure system on the left side of heart compared with the right side accounts for why the mitral and aortic valves are more affected than the pulmonary and tricuspid [1,21].

Previously, our research group studied these defects, considering the importance of lung irradiation and lung volume size in predicting the RVD toxicity risk. The influence of the left lung irradiation on radiation-induced lung fibrosis was also investigated [7,8]. Additionally, lung toxicity has very recently been correlated to lung volume size [22]. These results obtained in a clinical setting are consistent with those obtained from the Groningen group in animal studies [3,4]. Unfortunately, the patho-physiological mechanisms of heart-lung

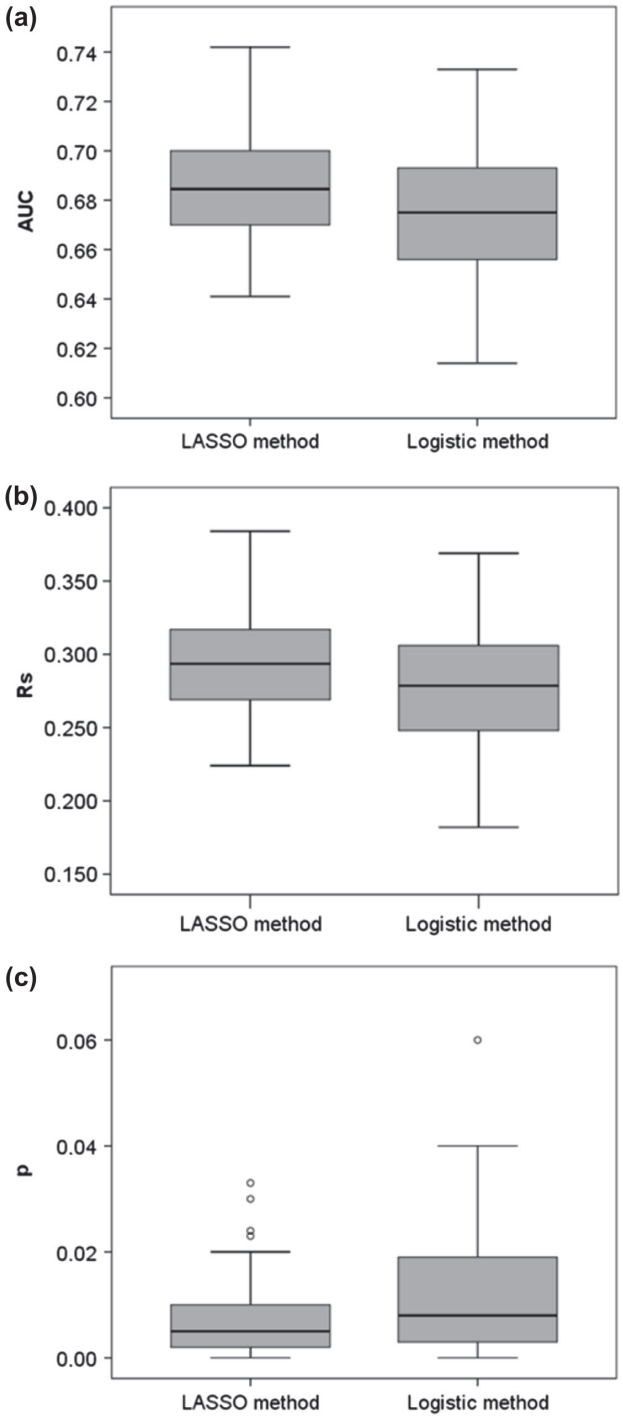

Figure 2. Comparison of prediction performance for LASSO logistic regression method and logistic regression method: (a) box plot for AUC; (b) box plot for Rs; (c) box plot for p-values.

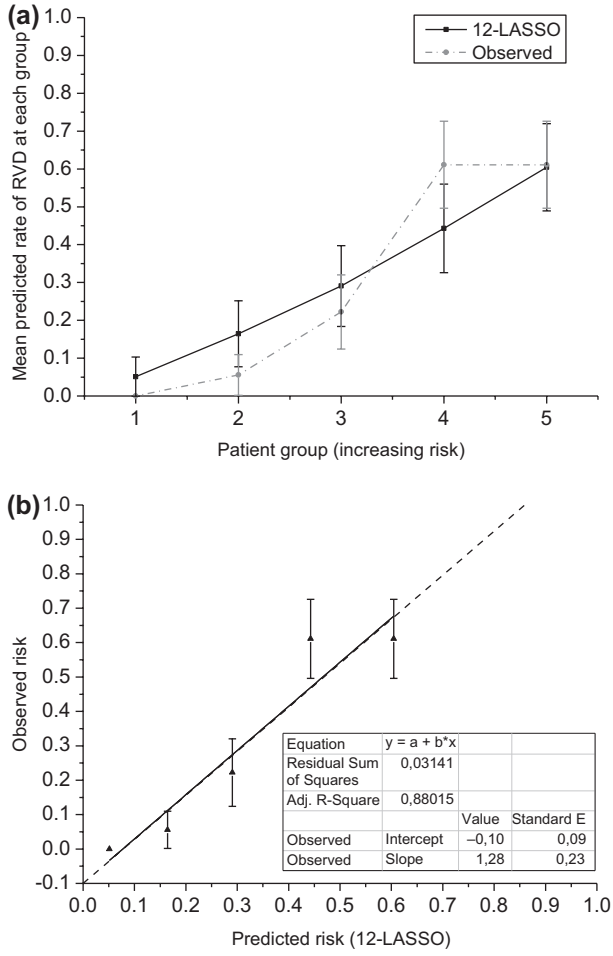

Figure 3. Mean predicted rates of radiation-induced valvular defects in binned groups (a) and calibration plot between the observed and predicted rates (b). The patients were binned into 5 groups based on the observed toxicity, with 1 being the lowest toxicity group and 5 being with the highest, with the equal number of patients (18) in each group.

interaction in the evolution of late toxicity after thoracic irradiation are still uncertain.

In the present study, our basic idea is to link all these results to perform a "clinical" radiation biology study through the development of robust predictive models. Analyzing an extended cohort of HL survivors including 90 patients (compared to 56 previously reported on [8]) with an extended follow-up, we found an RVD incidence of 30% with a left-side predominance compared with the right side (27.8% vs. 15.6%). Considering the age distribution of our patient cohort (Supplementary Table I available online at: http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1016624) the natural incidence of low to moderate valvular defects is expected to be lower than 2% [23]. Overall, the obtained results are in agreement with the existing literature reporting on valvular lesions in irradiated patients [19,20].

In the present study, thanks to an extended dataset and to an improved method for variable selection, namely LASSO, more variables have been taken into consideration. Applying the LASSO method with 10-fold CV, multiple optimized models were obtained and in each model, different variables were selected. Among the 26 candidate variables including heart along with left and right lung

Table II. Prognostic factors selected for 12- and 8-variables LASSO models, corresponding standardized regression coefficients β_i, odds ratios (OR) and transformed coefficients β_i^T expressed as the inverse measurement unit of the associated variable. Performance measures for each model are reported. For each variable, the mean and standard deviation (SD) used for standardization are indicated.

		12-LASSO			8-LASSO		
Variable	Mean ± SD	β_i	OR	β_i^T*	β_i	OR	β_i^T*
D10 heart (Gy)	24.7 ± 11.2	0.056	1.058	0.005	–	–	–
Dmax heart (Gy)	26.9 ± 11.0	0.083	1.087	0.008	–	–	–
AV30 heart (cm^3)	130.6 ± 174.2	0.248	1.281	0.001	0.149	1.161	0.001
Heart volume (cm^3)	558.0 ± 122.5	0.673	1.960	0.004	0.240	1.271	0.001
D15 left lung (Gy)	26.4 ± 10.0	0.320	1.377	0.032	0.220	1.246	0.022
D25 left lung (Gy)	19.7 ± 11.1	0.121	1.129	0.011	0.021	1.021	0.002
D95 left lung (Gy)	0.6 ± 0.6	0.061	1.063	0.102	0.138	1.148	0.230
Dmax left lung (Gy)	31.3 ± 7.3	0.299	1.349	0.041	–	–	–
Left lung volume (cm^3)	1278.2 ± 381.3	−0.792	0.453	−0.002	−0.165	0.848	−0.0004
D25 right lung (Gy)	20.4 ± 11.8	0.174	1.190	0.015	0.002	1.002	0.170
Dmax right lung (Gy)	31.4 ± 7.0	0.043	1.044	0.006	–	–	–
Dmax lungs (Gy)	33.1 ± 3.1	0.160	1.174	0.052	0.150	1.162	0.048
Constant		−1.150		−6.160	−0.906		−5.770
Performance measure							
Rs		0.55			0.51		
AUC (95% CI)		0.84 (0.77–0.93)			0.82 (0.73–0.91)		
Discrimination value		0.29			0.29		
Calibration slope		1.28 ± 0.23			2.15 ± 0.21		
Calibration intercept		−0.10 ± 0.09			−0.35 ± 0.07		

AUC, area under the ROC curve; AVx, absolute volume receiving x Gy; CI, confidence interval; Dx, the minimum dose to x% highest dose volume; Dmax, maximum dose; Rs, Spearman's correlation coefficient; Vx, percentage volume receiving x Gy.

*Performed transformation: $\beta_i^T = \frac{\beta_i}{SD_{xi}}$ and $\beta_0^T = \beta_0 - \sum \frac{\beta_i \, mean_{xi}}{SD_{xi}}$

dosimetric and clinical parameters, only 14 prognostic factors were found more than 100 times of 500 different models and only nine were selected with an occurrence higher than 50% (Figure 1b), thus representing more important predictive variables. Of these, most were related to left lung (D15, D25, D95, Dmax, volume) or lungs (Dmax) irradiation. The obtained results on the left lung were not related to an asymmetric irradiation of the different lungs. The only heart-related variables with high occurrence were the absolute volume of the heart receiving more than 30 Gy, heart Dmax, and the heart volume. Interestingly, the heart volume and the volume of left lung were both selected 481 times (96.2%). The right lung volume was selected only six times. The RVD risk increases with larger volume of the heart while it decreases when left lung volume size increases (Table II).

Although the LASSO has been successfully used in many practical problems, it has some limitations: 1) in the case of m (variables) $> n$ (patients), the LASSO selects at most n variables. However, this is not the case in our study ($n = 90$ and $m = 26$); 2) when there is a group of variables among which there is high correlation, the LASSO tends to select only one variable or a few variables (depending on the λ value) from the highly correlated variables [24]. A main goal of our study is to build a better model to predict the risk of RVD rather than to identify highly correlated variables. Therefore, the limitation is not a main issue in our study. It is likely that some variables in the same organ (lung or heart) are highly correlated. In our analysis, it was found that some lung and heart variables appeared together in LASSO models, implying that lung is associated with the development of RVD.

Concerning model uncertainty assessment, using 10-fold CV test, the performance of LASSO method was slightly better and most importantly it showed a smaller variability than the logistic regression with stepwise selection, with narrower box plots of AUC, Rs and p (Figure 2a-c).

We have to keep in mind that the adopted approach produces multiple NTCP models for common feature selection and for model statistical confidence evaluation, instead of a single best NTCP model. However, for practical purpose we also generated a final LASSO model starting from the 14 frequently selected variables. Overall, the final 12-variable LASSO model showed a good performance and discrimination capability (AUC = 0.844) as well as a high agreement between observed outcomes and predictions (Figure 3a and b, adjusted $R^2 = 0.88$).

This study is the first to suggest the importance of jointly considering left lung irradiation and left lung volume size in the prediction of subclinical radiation-related heart disease resulting in RVD. However, we cannot distinguish if these parameters are directly related to the analyzed outcome or carry information about a hidden parameter that is related to the outcome. Further external validation will be necessary for testing the clinical validity of this result.

Of note, this is not the first study to suggest that co-irradiation of the lung may influence heart function and vice versa [3,4,25]. The results of the present work represent a further confirmation that many variables are necessary to generate a reliable prediction tool able to consider a mechanism of multiple organ impairment, mainly when considering the cardio-pulmonary system.

As a whole, the obtained results pave the way to further investigations in order to clarify the still unclear role of lung and more specifically of left lung in the development of radiation-induced valve defects and in order to implement an effective individualized risk prediction.

Acknowledgments

The authors thank Dr. Marco Comerci for his technical support. This work was partially supported by grants from the Italian Ministry for Education, University and Research (MIUR) in the framework of FIRB (RBFR10Q0PT_001 "DROPS" and RBNE08YFN3 "MERIT").

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Jaworski C, Mariani JA, Wheeler G, Kaye DM. Cardiac complications of thoracic irradiation. J Am Coll Cardiol 2013;61:2319–28.

[2] Lancellotti P, Nkomo VT, Badano LP, Bergler-Klein J, Bogaert J, Davin L, et al. Expert consensus for multi-modality imaging evaluation of cardiovascular complications of radiotherapy in adults: A report from the European Association of Cardiovascular Imaging and the American Society of Echocardiography. Eur Heart J Cardiovasc Imaging 2013;14:721–40.

[3] Ghobadi G, van der Veen S, Bartelds B, de Boer RA, Dickinson MG, de Jong JR, et al. Physiological interaction of heart and lung in thoracic irradiation. Int J Radiat Oncol Biol Phys 2012;84:e639–46.

[4] van Luijk P, Faber H, Meertens H, Schippers JM, Langendijk JA, Brandenburg S, et al. The impact of heart irradiation on dose-volume effects in the rat lung. Int J Radiat Oncol Biol Phys 2007;69:552–9.

[5] Andratschke N, Maurer J, Molls M, Trott KR. Late radiation-induced heart disease after radiotherapy. Clinical importance, radiobiological mechanisms and strategies of prevention. Radiother Oncol 2011;100:160–6.

[6] Cella L, Palma G, Deasy JO, Oh JH, Liuzzi R, D'Avino V, et al. Complication probability models for radiation-induced heart valvular dysfunction: Do heart-lung interactions play a role? PloS One 2014;9:e111753.

[7] Cella L, Liuzzi R, D'Avino V, Conson M, Di Biase A, Picardi M, et al. Pulmonary damage in Hodgkin's lymphoma patients treated with sequential chemo-radiotherapy: Predictors of radiation-induced lung injury. Acta Oncol 2013.

[8] Cella L, Liuzzi R, Conson M, D'Avino V, Salvatore M, Pacelli R. Multivariate normal tissue complication probability modeling of heart valve dysfunction in Hodgkin lymphoma survivors. Int J Radiat Oncol Biol Phys 2013; 87:304–10.

[9] Lee TF, Chao PJ, Ting HM, Chang L, Huang YJ, Wu JM, et al. Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of xerostomia after intensity-modulated radiotherapy for head and neck cancer. PloS One 2014;9:e89700.

[10] Xu CJ, van der Schaaf A, Schilstra C, Langendijk JA, van't Veld AA. Impact of statistical learning methods on the predictive power of multivariate normal tissue complication probability models. Int J Radiat Oncol Biol Phys 2012; 82:e677–84.

[11] Xu CJ, van der Schaaf A, Van't Veld AA, Langendijk JA, Schilstra C. Statistical validation of normal tissue complication probability models. Int J Radiat Oncol Biol Phys 2012;84:e123–9.

[12] Feng M, Moran JM, Koelling T, Chughtai A, Chan JL, Freedman L, et al. Development and validation of a heart atlas to study cardiac exposure to radiation following treatment for breast cancer. Int J Radiat Oncol Biol Phys 2011;79:10–8.

[13] Kong FM, Ritter T, Quint DJ, Senan S, Gaspar LE, Komaki RU, et al. Consideration of dose limits for organs at risk of thoracic radiotherapy: Atlas for lung, proximal bronchial tree, esophagus, spinal cord, ribs, and brachial plexus. Int J Radiat Oncol Biol Phys 2011;81:1442–57.

[14] Deasy JO, Blanco AI, Clark VH. CERR: A computational environment for radiotherapy research. Med Phys 2003; 30:979–85.

[15] El Naqa I, Suneja G, Lindsay PE, Hope AJ, Alaly JR, Vicic M, et al. Dose response explorer: An integrated open-source tool for exploring and modelling radiotherapy dose-volume outcome relationships. Phys Med Biol 2006;51:5719–35.

[16] Aziz NM, Rowland JH. Trends and advances in cancer survivorship research: Challenge and opportunity. Semin Radiat Oncol 2003;13: 248–66.

[17] Veinot JP, Edwards WD. Pathology of radiation-induced heart disease: A surgical and autopsy study of 27 cases. Hum Pathol 1996;27:766–73.

[18] Cella L, Liuzzi R, Conson M, Torre G, Caterino M, De Rosa N, et al. Dosimetric predictors of asymptomatic heart valvular dysfunction following mediastinal irradiation for Hodgkin's lymphoma. Radiother Oncol 2011;101:316–21.

[19] Heidenreich PA, Hancock SL, Lee BK, Mariscal CS, Schnittger I. Asymptomatic cardiac disease following mediastinal irradiation. J Am Coll Cardiol 2003;42:743–9.

[20] Wethal T, Lund MB, Edvardsen T, Fossa SD, Pripp AH, Holte H, et al. Valvular dysfunction and left ventricular changes in Hodgkin's lymphoma survivors. A longitudinal study. Br J Cancer 2009;101:575–81.

[21] Adams MJ, Lipshultz SE, Schwartz C, Fajardo LF, Coen V, Constine LS. Radiation-associated cardiovascular disease:

Manifestations and management. Semin Radiat Oncol 2003;13:346–56.
[22] Briere T, Tucker S, Liao Z, Martel M. Small lung size increases the risk of radiation pneumonitis in lung cancer patients. Med Phys 2013;40:453.
[23] Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: A population-based study. Lancet 2006;368: 1005–11.
[24] Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Statist Soc B 2005;67:301–7.
[25] Huang EX, Hope AJ, Lindsay PE, Trovo M, El Naqa I, Deasy JO, et al. Heart irradiation as a risk factor for radiation pneumonitis. Acta Oncol 2011;50:51–60.

Supplementary material available online

Supplementary Figure 1 and 3, Table I and III, available online at: http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1016624.