DEVELOPMENT OF A MULTICOMPONENT PREDICTION MODEL FOR ACUTE ESOPHAGITIS IN LUNG CANCER PATIENTS RECEIVING CHEMORADIOTHERAPY

Int. J. Radiation Oncology Biol. Phys., Vol. 81, No. 2, pp. 537–544, 2011

0360-3016/$ - see front matter

doi:10.1016/j.ijrobp.2011.03.012

BIOLOGY CONTRIBUTION

DEVELOPMENT OF A MULTICOMPONENT PREDICTION MODEL FOR ACUTE ESOPHAGITIS IN LUNG CANCER PATIENTS RECEIVING CHEMORADIOTHERAPY

Kim De Ruyck, Ph.D.,* Nick Sabbe, M.Sc.,† Cary Oberije, Ph.D.,‡ Katrien Vandecasteele, M.D.,§ Olivier Thas, Ph.D.,† Dirk De Ruysscher, M.D., Ph.D.,‡ Phillipe Lambin, M.D., Ph.D.,‡ Jan Van Meerbeeck, M.D., Ph.D.,‖ Wilfried De Neve, M.D., Ph.D.,§ and Hubert Thierens, Ph.D.*

Departments of *Basic Medical Sciences and †Applied Mathematics, Biometrics and Process Control, Ghent University, Ghent, Belgium; ‡Department of Radiation Oncology (MAASTRO Clinic), Research Institute of Growth and Development, Maastricht University Medical Center, Maastricht, The Netherlands; and Departments of §Radiation Oncology and ‖Respiratory Medicine, Ghent University Hospital, Ghent, Belgium

Purpose: To construct a model for the prediction of acute esophagitis in lung cancer patients receiving chemoradiotherapy by combining clinical data, treatment parameters, and genotyping profile.

Patients and Methods: Data were available for 273 lung cancer patients treated with curative chemoradiotherapy. Clinical data included gender, age, World Health Organization performance score, nicotine use, diabetes, chronic disease, tumor type, tumor stage, lymph node stage, tumor location, and medical center. Treatment parameters included chemotherapy, surgery, radiotherapy technique, tumor dose, mean fractionation size, mean and maximal esophageal dose, and overall treatment time. A total of 332 genetic polymorphisms were considered in 112 candidate genes. The predicting model was achieved by lasso logistic regression for predictor selection, followed by classic logistic regression for unbiased estimation of the coefficients. Performance of the model was expressed as the area under the curve of the receiver operating characteristic and as the false-negative rate in the optimal point on the receiver operating characteristic curve.

Results: A total of 110 patients (40%) developed acute esophagitis Grade ≥2 (Common Terminology Criteria for Adverse Events v3.0). The final model contained chemotherapy treatment, lymph node stage, mean esophageal dose, gender, overall treatment time, radiotherapy technique, rs2302535 (*EGFR*), rs16930129 (*ENG*), rs1131877 (*TRAF3*), and rs2230528 (*ITGB2*). The area under the curve was 0.87, and the false-negative rate was 16%.

Conclusion: Prediction of acute esophagitis can be improved by combining clinical, treatment, and genetic factors. A multicomponent prediction model for acute esophagitis with a sensitivity of 84% was constructed with two clinical parameters, four treatment parameters, and four genetic polymorphisms. © 2011 Elsevier Inc.

Prediction, Esophagitis, Radiotherapy, Genetic polymorphisms, Lasso logistic regression.

INTRODUCTION

Lung cancer has the highest incidence and mortality rate of all cancers in Western countries, with most patients presenting with advanced-stage disease at the time of diagnosis (1). The standard of care for locally advanced non–small-cell lung cancer is concurrent chemoradiotherapy. Treatment success is, however, still constrained by poor local control and posttherapy toxicity as acute esophagitis (2, 3).

Numerous studies have attempted to define clinical and dosimetric predictors of radiation-induced esophagitis (4–21). Factors found to correlate with important acute esophagitis include concurrent chemoradiotherapy, lymphatic status, and a number of dose–volumetric parameters. Because the results varied considerably across different institutions, their clinical usefulness remains restricted. Simultaneously, significant research efforts have been made to link genetic polymorphisms in selected genes to radiation-induced toxicity (22–30). The majority of these radiogenetics studies considered patients with breast or prostate cancer, whereas only a few research groups reported on radiation-induced toxicity after radiotherapy for lung cancer (28–31). Because of the limited genetic studies performed for esophagitis and the inconsistent outcomes for other radiation-induced toxicities, genetic biomarkers are not yet implementable in the clinic.

Reprint requests to: Kim De Ruyck, Ph.D., Ghent University, Department of Basic Medical Sciences, Proeftuinstraat 86, B-9000 Ghent, Belgium. Tel: (+32)-926-46656; Fax: (+32) 926-46696; E-mail: kim.deruyck@UGent.be

Supported by Grant No. 365D6706 from the Stichting tegen Kanker and Grant No. 09/PDO/061 from the Fonds voor Wetenschappelijk Onderzoek.

Conflicts of interest: none.

Acknowledgments—The authors thank all the study participants and data managers, Dr. Bart Loeys, and Virginie de Gelder.

Received Jan 25, 2011, Accepted for publication March 8, 2011.

Over the years it has become clear that the effects of dose distribution, clinical parameters, and individual genetic variation on radiation-induced toxicity may not be evaluated separately. As a result, studies are emerging that correct for dosimetric and patient-related risk factors when trying to link genetic polymorphisms with a clinical phenotype after therapy (29, 31–34). However, a model to predict susceptibility to radiation in individual patients is still unavailable. Therefore, we aimed at constructing a predictive algorithm for acute esophagitis by combining clinical data, treatment parameters, and genotypic information. This would enable us to individualize patient treatment.

PATIENTS AND METHODS

Study population

Two hundred eighty-nine lung cancer patients treated with curative radiotherapy between February 2004 and August 2009 were enrolled. Of these, 273 were suited to perform the study (Fig. 1). A total of 213 patients were recruited from the MAASTRO Clinic and 60 patients from the Ghent University Hospital. Clinical data and treatment details are presented in Table 1. The majority of patients were treated with three-dimensional conformal radiotherapy (3D-CRT) as opposed to intensity-modulated radiotherapy (IMRT). The median tumor dose was 60 Gy at 1.5–2.69 Gy per fraction (2 patients received 7.5 Gy per fraction). The details of the different radiotherapy treatment regimens can be found in Appendix E1 (available online). For the MAASTRO Clinic patients, the dosimetric parameters were calculated using a commercial radiotherapy treatment planning system (XiO; Computerized Medical Systems, St. Louis, MO). The dosimetric parameters for the Ghent patients were calculated using an in-house-developed planning system, with a final dose computation using a commercial radiotherapy planning system (Pinnacle; Philips Medical Systems, Best, The Netherlands). In both centers, the esophagus was delineated using the external esophageal contour from the cricoid cartilage to the gastroesophageal junction. Esophageal toxicity was scored using the Common Terminology Criteria for Adverse Events scale version 3.0 (35). Acute esophagitis was defined as dysphagia Grade ≥2 at any time during or at maximum 3 months after radiotherapy treatment. Genomic DNA was obtained from fresh blood (Ghent samples) or frozen buffy coat (MAASTRO samples) using the Puregene genomic DNA purification kit (Gentra Systems, Minneapolis, MN). The study was approved by the ethics committees of both centers, and all study participants provided written informed consent. The MAASTRO Clinic cohort study was filed at clinicaltrials.gov (no. NCT01084785).

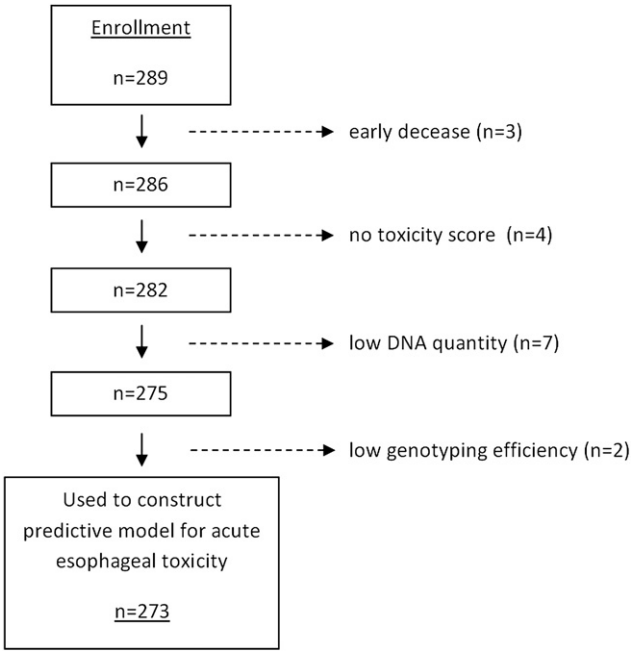

Fig. Study flowchart.

Selection of candidate genes and genetic variations

On the basis of a literature search, a total of 112 candidate genes belonging to the following categories were considered: early response genes ($n = 15$), cytokines and growth factors ($n = 14$), signal transduction of cytokines and growth factors ($n = 53$), adhesion molecules ($n = 9$), extracellular matrix genes ($n = 13$), and others ($n = 8$). Single nucleotide polymorphisms (SNPs) were selected after functional tagging based on evolutionary conserved regions (ECRs) using ECR Browser (36). The 5′ flanking region of each gene (5 kb) was included to thoroughly examine any possible regulatory regions. Conserved regions with a minimal length of 100 bp with minimal 80% equality over four species (mouse, rat, rhesus monkey, and human) were considered. The criteria for picking SNPs in the ECRs were as follows: minor allele frequencies in populations of Caucasian ethnicity >5%, Illumina SNP score >0.60 (proprietary score used to determine the overall success of the assay), and no linkage ($r^2 > 0.80$) with other SNPs. Less-stringent criteria were used for cytokines, their receptors, and cell adhesion molecules (ECRs of 75% equality over three species). The selection was expanded with a number of SNPs with a proven biological function from the literature. Finally, 384 SNPs were retained (Table E1).

Genotyping and quality control

The DNA quality and quantity were checked before genotyping, and 7 individuals were omitted from genotyping because of low DNA quantity. Genotyping was performed using the Illumina Goldengate technology (DNA Vision, Charleroi, Belgium). Upon completion of genotyping, quality control (QC) processes were run to guarantee the accuracy of the genotyped dataset. Two individuals were dropped on the basis of low genotyping efficiency (<90%). Single nucleotide polymorphisms were eliminated on the basis of low reliability of cluster separation, low signal intensity, low call frequency (<75%), absence of a minor allele in the dataset, and deviation from Hardy-Weinberg equilibrium ($p < 0.0001$). After application of stringent QC, approximately 3% of the samples and approximately 13% of the SNPs were eliminated. Overall, QC was completed yielding 273 individuals and 332 SNPs (Table E1).

Statistics

The predicting model was achieved in two steps. First, lasso logistic regression was applied to the full dataset, for which the lasso parameter was chosen so as to maximize the area under the curve (AUC) of the receiver operator characteristic (ROC) curve. The latter was estimated from 10-fold cross-validation (37). Second, the implied set of predictors was passed to classic logistic regression for unbiased estimation of the coefficients.

Lasso regression shrinks the coefficient estimates toward zero, with the degree of shrinkage depending on an additional parameter, lambda (λ). In this way, coefficient estimates can be forced to be exactly zero, thereby effectively eliminating a number of variables.

Table 1. Clinical and treatment parameters for esophageal toxicity

Parameter	All (n = 273)	Grade 0–1 (n = 163)	Grade 2+ (n = 110)
Clinical parameters			
Gender (n)			
Men	200 (73.3)	136 (83.4)	64 (58.2)
Woman	73 (26.7)	27 (16.6)	46 (41.8)
Age (y)			
Median	66.6	68.5	61.5
Range	42.8–87.0	44.0–87.0	42.8–86.0
WHO performance score (n)			
0	88 (32.2)	53 (32.5)	35 (31.8)
1	144 (52.7)	84 (51.5)	60 (54.5)
≥2	39 (14.3)	25 (15.3)	14 (12.7)
Missing	2 (0.7)	1 (0.6)	1 (0.9)
Nicotine use (n)			
Never	7 (2.6)	4 (2.5)	3 (2.7)
Current	93 (34.1)	48 (29.4)	45 (40.9)
Former	165 (60.4)	107 (65.6)	58 (52.7)
Missing	8 (2.9)	4 (2.5)	4 (3.6)
Diabetes (n)			
Yes	23 (8.4)	18 (11.0)	5 (4.5)
No	231 (84.6)	141 (86.5)	90 (81.8)
Missing	19 (7.0)	4 (2.5)	15 (13.6)
Chronic disease* (n)			
Yes	7 (2.6)	3 (1.8)	4 (3.6)
No	250 (91.6)	154 (94.5)	96 (87.3)
Missing	16 (5.9)	6 (3.7)	10 (9.1)
Tumor type (n)			
NSCLC	209 (76.6)	135 (82.8)	74 (67.3)
SCLC	50 (18.3)	17 (10.4)	33 (30.0)
Missing	14 (5.1)	11 (6.7)	3 (2.7)
Tumor stage (n)			
I	40 (14.7)	36 (22.1)	4 (3.6)
II	27 (9.9)	19 (11.7)	8 (7.3)
IIIa	88 (32.2)	59 (36.2)	29 (26.4)
IIIb	102 (37.4)	43 (26.4)	59 (53.6)
IV	5 (1.8)	1 (0.6)	4 (3.6)
Missing	11 (4.0)	5 (3.1)	6 (5.5)
Lymph node stage (n)			
N0	83 (30.4)	67 (41.1)	16 (14.5)
N1	20 (7.3)	15 (9.2)	5 (4.5)
N2	95 (34.8)	58 (35.6)	37 (33.6)
N3	64 (23.4)	18 (11.0)	46 (41.8)
Missing	11 (4.0)	5 (3.1)	6 (5.5)
Tumor location (n)			
Lower lobe	61 (22.3)	34 (20.9)	27 (24.5)
Other	200 (73.3)	123 (75.5)	77 (70.0)
Missing	12 (4.4)	6 (3.7)	6 (5.5)
Medical center (n)			
Ghent University Hospital	60 (22.0)	46 (28.2)	14 (12.7)
MAASTRO Clinic	213 (78.0)	117 (71.8)	96 (87.3)
Treatment parameters			
Chemotherapy (n)			
No	69 (25.3)	55 (33.7)	14 (12.7)
Sequential	120 (44.0)	83 (50.9)	37 (33.6)
Concurrent	84 (30.8)	25 (15.3)	59 (53.6)
Surgery (n)			
Yes	35 (12.8)	25 (15.3)	10 (9.1)
No	238 (87.2)	138 (84.7)	100 (90.9)
Radiotherapy technique (n)			
3D-CRT	224 (82.1)	124 (76.1)	100 (90.9)
IMRT	49 (17.9)	39 (23.9)	10 (9.1)
Tumor dose (Gy)			
Median	60.0	64.0	54.0
Range	40.0–79.2	40.0–79.2	45.0–79.2

(*Continued*)

Table 1. Clinical and treatment parameters for esophageal toxicity (*Continued*)

Parameter	All (n = 273)	Grade 0–1 (n = 163)	Grade 2+ (n = 110)
Mean fractionation size (Gy)			
Median	1.8	1.8	1.8
Range	1.5–7.5	1.5–7.5	1.5–2.5
Mean esophageal dose (Gy)			
Median	21.7	16.3	26.1
Range	0.4–45.7	0.4–41.9	3.5–45.7
Missing	22	13	9
Maximum esophageal dose (Gy)			
Median	56.3	56.0	56.5
Range	1.1–84.2	1.1–81.8	22.8–84.2
Missing	22	14	8
Overall treatment time (d)			
Median	31.0	36.0	25.0
Range	17.0–62.0	17.0–62.0	19.0–61.0

Abbreviations: WHO = World Health Organization; NSCLC = non–small-cell lung cancer; 3D-CRT = three-dimensional conformal radiotherapy; IMRT = intensity-modulated radiotherapy.

Data in parentheses are percentages.

* Gastroesophageal reflux disease, Barrett's syndrome, previous larynx carcinoma.

By varying λ from 0 to 1, and, for each value of λ, evaluating a measure of predictive value (AUC) by cross-validation, the optimal λ is found. This choice of λ corresponds to a fitted lasso regression model for which a limited set of variables provides an acceptable level of predictive stability, whereas adding more variables does not increase the predictive value. As a result, a set of predictors is obtained for the selected λ value (*i.e.*, the lasso step acts as a model selection procedure). A consequence of the lasso technique is that all coefficient estimates tend to be downwardly biased. Therefore in a second step, a classic non-lasso logistic regression was performed to find unbiased estimates of the coefficients of the selected set of predictors. Again, the predictive value was evaluated through 10-fold cross-validation. The problem of missing values was surmounted by applying multiple imputation (38–40). For each step described above, 500 full datasets were imputed using predictive mean matching for continuous variables and polytomous logistic regression for categoric variables. The analysis was applied to each of these full datasets, after which the results were averaged and the variances of the coefficient estimates were corrected for the variability added by the imputation. Because the AUC reflects the full spectrum of misclassification, it is only partially suited in this study, because our interest goes particularly to the part of the ROC curve where false negatives are unlikely. To this end, a criterion for a balanced "optimal point" on the ROC curve was

Table 2. Outcome of modeling analyses

Parameter	Occurrence* (%)	Two-SNP model	Four-SNP model	Non-SNP model
Model predictors, coefficient (95% CI)				
(Intercept)		−4.29 (−6.45, −2.14)	−4.07 (−6.39, −1.76)	−3.71 (−5.78, −1.65)
Concurrent chemotherapy	100	1.70 (0.10, 2.40)	1.85 (1.10, 2.61)	1.74 (1.06, 2.42)
Lymph node stage N3	100	1.05 (0.21, 1.88)	1.24 (0.35, 2.13)	1.02 (0.23, 1.81)
Mean esophageal dose	100	0.09 (0.05, 0.13)	0.09 (0.05, 0.14)	0.08 (0.04, 0.12)
Gender female	100	0.97 (0.26, 1.68)	1.04 (0.30, 1.79)	1.02 (0.34, 1.71)
Overall treatment time	100	−0.02 (−0.06, 0.01)	−0.02 (−0.06, 0.02)	−0.02 (−0.05, 0.02)
RT technique 3D-CRT	78.8	0.99 (−0.08, 2.06)	1.05 (−0.06, 2.16)	1.10 (0.05, 2.15)
rs2302535CC *(EGFR)*	59.4	1.67 (0.31, 3.03)	1.82 (0.42, 3.21)	—
rs16930129CC *(ENG)*	56.4	0.97 (0.19, 1.75)	0.91 (0.11, 1.72)	—
rs1131877AG *(TRAF3)*	47.4	—	−1.14 (−1.89, −0.38)	—
rs2230528GA *(ITGB2)*	44.4	—	−0.75 (−1.47, −0.03)	—
Model characteristics, % (95% CI)				
Probability threshold†	—	30.7 (21.9, 39.5)	32.3 (23.9, 40.6)	37.0 (29.3, 44.9)
Misclassification rate†	—	23.6 (21.5, 25.6)	21.2 (19.2, 23.3)	23.9 (21.0, 26.8)
False positive rate†	—	28.2 (20.8, 35.5)	24.7 (18.9, 30.6)	24.7 (16.6, 32.5)
False negative rate†	—	16.7 (8.8, 24.8)	16.0 (9.9, 22.2)	22.8 (16.6, 29.2)
AUC	—	85.2 (83.9, 86.5)	86.7 (85.4, 88.0)	84.1 (82.8, 85.4)

Abbreviations: SNP = single nucleotide polymorphism; CI = confidence interval; RT = radiotherapy; 3D-CRT = three-dimensional conformal RT; AUC = area under the curve.

Model expression: $\ln p/(1-p) = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \beta_3 \times x_3 + \beta_4 \times x_4 + \beta_5 \times x_5 + \beta_6 \times x_6 + \ldots$ with p = probability to develop acute esophagitis, β = coefficient, x = variable/predictor, and $e^{\beta 1}$ = increase in odds for predictor 1 (in the case of continuous variables: for an increase of the predictor by 1 unit) when the other predictors remain unchanged.

* Over all imputations and partial fits, how often is the predictor included in the best selected model?

† In the optimal point on the receiver operator characteristic curve.

used: false negatives received a relative weight of 1.5 while minimizing the distance to the perfect point of the ROC curve (top left). This resulted in a threshold for the prediction decision rule, as well as estimates for the false-negative and false-positive rates. The confidence intervals (CI) of the coefficients are the ordinary logistic regression CIs, corrected for multiple imputation. All analyses were performed in R by applying the R package glmnet (41) for lasso regression and the R package MICE (42) for multiple imputation.

RESULTS

In total, 351 variables were available for model building. The 19 clinical/treatment parameters are listed in Table 1, and the 332 genetic polymorphisms are available in Table E1. After the variable selection procedure including all variables that occur in more than 50% of the imputed datasets, the model contained chemotherapy, lymph node stage, mean esophageal dose (MED), gender, overall treatment time (OTT), radiation technique, rs2302535, and rs16930129 (Table 2). Lowering the variable selection threshold to 40% resulted in the addition of the rs1131877 and rs2230528 SNPs to the model (Table 2). Estimation of the predictor coefficients showed that, for both approaches, the most powerful risk predictors were concurrent chemotherapy and the rs2302535CC genotype. The AUCs of the two-SNPs model and the four-SNPs model were 0.85 and 0.87, respectively. The false-negative rate in the optimal point of the ROC curve was slightly smaller for the model including four SNPs (16.0% vs. 16.7%). Analogous to the representation of the clinical/treatment variables in Table 1, the data for the four SNPs included in the models are presented in Table 3. Predictor coefficients and model characteristics were also estimated for a model excluding genetic information. This non-SNP model had an AUC of 0.84 and a false-negative rate in the optimal point of the ROC curve of 22.8% (Table 2). The 95% CIs for the predictor coefficients are rather broad, whereas those for the measures of predictive value are relatively small. This indicates that the predictive accuracy using the given predictors and coefficients is stable.

Table 3. Data for SNPs included in models

SNP	All (n = 273)	Grade 0–1 (n = 163)	Grade 2+ (n = 110)
rs2302535 (*EGFR*)			
AA	133 (48.7)	87 (53.4)	46 (41.8)
AC	121 (44.3)	71 (43.6)	50 (45.5)
CC	19 (7.0)	5 (3.1)	14 (12.7)
Missing	0	0	0
rs16930129 (*ENG*)			
CC	207 (75.8)	114 (69.9)	93 (84.5)
CT	64 (23.4)	47 (28.8)	17 (15.5)
TT	2 (0.7)	2 (1.2)	0 (0.0)
Missing	0	0	0
rs1131877 (*TRAF3*)			
AA	164 (60.3)	92 (56.4)	72 (66.1)
AG	95 (34.9)	65 (39.9)	30 (27.5)
GG	13 (4.8)	6 (3.7)	7 (6.4)
Missing	1	0	1
rs2230528 (*ITGB2*)			
GG	151 (55.7)	82 (50.6)	69 (63.3)
GA	107 (39.5)	74 (45.7)	33 (30.3)
AA	13 (4.8)	6 (3.7)	7 (6.4)
Missing	2	1	1

Abbreviation as in Table 2.
Data in parentheses are genotype frequencies.

DISCUSSION

Acute esophagitis is one of the most important side effects of high-dose radiotherapy for lung cancer, especially when combined with concurrent chemotherapy (2, 43). Currently, the prediction of this side effect is based on dosimetric parameters of radiotherapy only. However, these parameters lack sensitivity and specificity to estimate patient-specific treatment outcome correctly. To increase their predictive value, additional parameters are required. Different association studies point to the importance of the individual genotypic profile in therapy toxicity. Therefore, this study combined clinical data, treatment parameters, and genetic variation to develop a predictive multicomponent model for acute esophagitis. Single nucleotide polymorphisms in multispecies conserved sequences of candidate genes were considered. Conserved sequences have remained similar across the millions of years of evolution and are believed to indicate regions of biological function (44). By using conservation to prioritize SNPs, the chances may be increased that SNPs impacting phenotype will actually be genotyped. Including genetic variation data to model building may add to the model's predictive value but also complicates the analysis owing to the strong increase in the number of variables. To get around this problem, lasso logistic regression was applied because this technique finds models involving the smallest number of parameters while preserving predictive value. At the same time, variable inclusion is based on predictive value (as opposed to statistical significance) and over-/underfitting is avoided. In the present study, two clinical (lymph node stage and gender), four treatment (chemotherapy, MED, OTT, and radiation technique), and four genetic (rs2302535, rs16930129, rs1131877, and rs2230528) parameters were found to be highly predictive for the development of Grade $\geq$2 acute esophageal postradiotherapy toxicity.

Concurrent chemotherapy treatment has been accepted as a risk factor for acute esophagitis. In our study, in patients treated without or with concurrent chemotherapy, Grade $\geq$2 acute esophagitis incidences were 27.0% and 70.2%, respectively. This increase in toxicity for concurrent chemotherapy has also been found in a high number of other studies (4, 5, 14, 16, 18–21). In parallel, published data demonstrate that the best-studied dosimetric parameters with high levels of association with acute esophagitis are MED, V_{20}, V_{30}, V_{40}, V_{45},V_{50}, and V_{55} (4–6, 8, 9).

The predicting quality of lymph node stage could be explained by the correlation between the extent of lymph node involvement and the irradiated volume and dose to the

esophagus, especially in patients with N3 disease (correlation coefficient with MED, 0.42). The median MED was 29.1 Gy for the N3 group, compared with 18.5 Gy for the N0/1/2 group. The incidence of Grade ≥2 acute esophagitis was 71.9% in the N3 arm, compared with 29.2% in the N0/1/2 arm. These results generally agree with a number of other studies that found an association of nodal stage N2/3 with Grade ≥2 acute esophagitis (5, 13–15). Although in previous studies gender was only exceptionally associated with acute esophagitis (4, 20), it was found to be a predictor in the present study. The incidence of Grade ≥2 acute esophagitis was 63.0% for women and 32.0% for men.

Because of the use of different radiotherapy treatment regimens, treatment times varied considerably between patients, and treatment duration could be included in the analysis. As expected according to radiobiologic principles, short treatment times were found to be predictive for increased acute esophageal toxicity. This is in line with some (4, 20) but not all studies (10, 13). Compared with IMRT, 3D-CRT was predictive for acute esophagitis in the present study. This may be explained by the shorter treatment times for patients irradiated with 3D-CRT rather than by the dose to the esophagus (correlation coefficients of −0.52 and −0.09, respectively). The median treatment time was 28 days for the 3D-CRT group and 47 days for the IMRT group, whereas the median MED was, respectively, 21.2 Gy and 26.1 Gy. The incidence rate in the 3D-CRT arm was 44.6%, compared with 20.4% in the IMRT arm. Alternatively, higher doses of radiation to parts of the esophagus outside of the planning target volumes, occurring more in 3D-CRT than in IMRT, may also explain the higher incidence of esophagitis in the 3D-CRT group.

The rs2302535 SNP is situated in intron 2 of the epidermal growth factor receptor gene (*EGFR*). Epidermal growth factor receptor plays an important role in cellular signaling and is central to human tumorigenesis. Because in cancer cells the activation of EGFR leads to a series of intracellular signals resulting in increased tumor cell growth and resistance to apoptosis, EGFR inhibitors have found their use in anticancer treatment (45). In this context, *EGFR* polymorphisms are being studied in relation to disease response and toxicity after EGFR inhibitor therapy (46, 47). The rs16930129 SNP (Leu69Leu) is located in exon 2 of the endoglin gene (*ENG*). Endoglin is a transforming growth factor (TGF)-β coreceptor mainly expressed in endothelial cells. It regulates cell proliferation and is important for endothelial cell survival and vessel repair. Mutations in endoglin have been linked to the vascular disease hereditary hemorrhagic telangiectasia, characterized by dilated capillaries (telangiectasia) that are prone to rupture (48). Radiation-induced vascular damage strongly resembles the symptoms of hereditary hemorrhagic telangiectasia patients, and animal studies showed an important role for endoglin in the development of radiation-induced normal tissue damage (49, 50). The rs1131877 SNP (Met129Thr) is situated in exon 3 of the tumor necrosis factor receptor-associated factor 3 gene (*TRAF3*). The encoded protein is a member of the TRAF family, which is important in activating multiple inflammatory and immune-related processes induced by cytokines such as tumor necrosis factor-α and interleukin-1 (51). The rs2230528 SNP (Gly273Gly) is part of exon 7 of the integrin β2 gene (*ITGB2*). Integrin β2 belongs to the integrin β-chain family proteins that may combine with α-chain family proteins to form β2 integrins. These β2 integrins are leukocyte cell adhesion molecules including LFA-1 and Mac-1, with an important role in cell-surface–mediated signaling and radiation-mediated inflammation. It has been found that radiation induces changes in the expression of various integrins, with important implications for tumor control and normal tissue toxicity (52).

Generally, of the four genetic variants that were found to be predictive for acute esophagitis, three are part of signal transduction genes of cytokines and growth factors, whereas one is localized in a gene coding for a cell adhesion molecule. Although the encoded proteins are involved in radiation-induced reactions, studies linking polymorphisms in these genes to radiation toxicity have not been performed. Haploview analysis showed that the *EGFR* rs2302535 SNP is tightly linked with two nearby polymorphisms and that the *TRAF3* rs1131877 SNP is strongly linked with a high number of intronic polymorphisms (mean maximum r^2 of 0.96 and 0.98, respectively). A high degree of linkage ($r^2 > 0.8$) with other polymorphisms was not found for the *ENG* and the *ITGB2* SNPs. At present, the functional role of any of these polymorphisms is unknown. When a classical association analysis was performed on the dataset of the present study, none of the 332 considered SNPs were significantly associated with Grade ≥2 acute esophagitis after correction for multiple testing with either Benjamini-Hochberg or Bonferroni tests. Only by omitting this correction, a number of polymorphisms, including the four retained in the prediction model, were significantly associated with esophagitis. The two *TGFβ1* SNPs (rs1800469/-509 and rs1800470/Leu10Pro), which were previously shown to be significantly associated with radiation-induced esophagitis or pneumonitis (30, 31), were not significantly associated with acute esophagitis in the present study and were also not retained in the model.

The final predictive model including 10 parameters estimates the risk of Grade ≥2 acute esophagitis with 84.0% sensitivity and 75.3% specificity. By adding genetic data to the clinical and treatment parameters, the sensitivity of the model increased from 77% to 84%. In our opinion, further progress is possible by genotyping a higher number of SNPs in a genome-wide approach. To evaluate the clinical usefulness of the biomarker, its predictive value was determined. With an incidence rate of 40% for Grade ≥2 acute esophagitis as observed in the patient population of the present study, the positive predictive value of the biomarker is 69.4%, and the negative predictive value is 87.4%. In other words, 7 in 10 patients predicted to suffer from acute esophagitis will actually develop this toxicity and will

benefit from therapy modification. Likewise, 9 in 10 patients will be correctly predicted as low-risk patients and will be suitable candidates for dose-escalation studies. In conclusion, the good performance of the model may be the basis of new prospective trials for treatment optimization in lung cancer patients treated with chemoradiation. Prospective validation in an independent patient cohort has started, to confirm this result.

REFERENCES

1. Jemal A, Siegel R, Xu J, *et al*. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277–300.
2. Auperin A, Le Pechoux C, Rolland E, *et al*. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J Clin Oncol* 2010;28: 2181–2190.
3. De Ruysscher D, Faivre-Finn C, Nestle U, *et al*. European Organisation for Research and Treatment of Cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. *J Clin Oncol* 2010;28: 5301–5310.
4. Dehing-Oberije C, De Ruysscher D, Petit S, *et al*. Development, external validation and clinical usefulness of a practical prediction model for radiation-induced dysphagia in lung cancer patients. *Radiother Oncol* 2010;97:455–461.
5. Zhang ZC, Xu J, Li BS, *et al*. Clinical and dosimetric risk factors of acute esophagitis in patients treated with 3-dimensional conformal radiotherapy for non-small-cell lung cancer. *Am J Clin Oncol* 2010;33:271–275.
6. Rose J, Rodrigues G, Yaremko B, *et al*. Systematic review of dose-volume parameters in the prediction of esophagitis in thoracic radiotherapy. *Radiother Oncol* 2009;91:282–287.
7. Watkins JM, Wahlquist AE, Shirai K, *et al*. Factors associated with severe acute esophagitis from hyperfractionated radiotherapy with concurrent chemotherapy for limited-stage small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2009;74:1108–1113.
8. Rodriguez N, Algara M, Foro P, *et al*. Predictors of acute esophagitis in lung cancer patients treated with concurrent three-dimensional conformal radiotherapy and chemotherapy. *Int J Radiat Oncol Biol Phys* 2009;73:810–817.
9. Topkan E, Yavuz MN, Onal C, *et al*. Prevention of acute radiation-induced esophagitis with glutamine in non-small cell lung cancer patients treated with radiotherapy: Evaluation of clinical and dosimetric parameters. *Lung Cancer* 2009;63: 393–399.
10. De Ruysscher D, Dehing C, Bremer RH, *et al*. Maximal neutropenia during chemotherapy and radiotherapy is significantly associated with the development of acute radiation-induced dysphagia in lung cancer patients. *Ann Oncol* 2007;18:909–916.
11. Wei X, Liu HH, Tucker SL, *et al*. Risk factors for acute esophagitis in non-small-cell lung cancer patients treated with concurrent chemotherapy and three-dimensional conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2006;66:100–107.
12. Takeda K, Nemoto K, Saito H, *et al*. Predictive factors for acute esophageal toxicity in thoracic radiotherapy. *Tohoku J Exp Med* 2006;208:299–306.
13. Ahn SJ, Kahn D, Zhou SM, *et al*. Dosimetric and clinical predictors for radiation-induced esophageal injury. *Int J Radiat Oncol Biol Phys* 2005;61:335–347.
14. Belderbos J, Heemsbergen W, Hoogeman M, *et al*. Acute esophageal toxicity in non-small cell lung cancer patients after high dose conformal radiotherapy. *Radiother Oncol* 2005;75: 157–164.
15. Chapet O, Kong FM, Lee JS, *et al*. Normal tissue complication probability modeling for acute esophagitis in patients treated with conformal radiation therapy for non-small cell lung cancer. *Radiother Oncol* 2005;77:176–181.
16. Kim TH, Cho KH, Pyo HR, *et al*. Dose-volumetric parameters of acute esophageal toxicity in patients with lung cancer treated with three-dimensional conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2005;62:995–1002.
17. Patel AB, Edelman MJ, Kwok Y, *et al*. Predictors of acute esophagitis in patients with non-small-cell lung carcinoma treated with concurrent chemotherapy and hyperfractionated radiotherapy followed by surgery. *Int J Radiat Oncol Biol Phys* 2004;60:1106–1112.
18. Bradley J, Deasy JO, Bentzen S, *et al*. Dosimetric correlates for acute esophagitis in patients treated with radiotherapy for lung carcinoma. *Int J Radiat Oncol Biol Phys* 2004;58:1106–1113.
19. Singh AK, Lockett MA, Bradley JD. Predictors of radiation-induced esophageal toxicity in patients with non-small-cell lung cancer treated with three-dimensional conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;55:337–341.
20. Werner-Wasik M, Pequignot E, Leeper D, *et al*. Predictors of severe esophagitis include use of concurrent chemotherapy, but not the length of irradiated esophagus: A multivariate analysis of patients with lung cancer treated with nonoperative therapy. *Int J Radiat Biol Oncol Phys* 2000;48:689–696.
21. Maguire PD, Sibley GS, Zhou SM, *et al*. Clinical and dosimetric predictors of radiation-induced esophageal toxicity. *Int J Radiat Oncol Biol Phys* 1999;45:97–103.
22. Andreassen CN, Alsner J. Genetic variants and normal tissue toxicity after radiotherapy: A systematic review. *Radiother Oncol* 2009;92:299–309.
23. Alsbeih G, Al-Harbi N, Al-Hadyan K, *et al*. Association between normal tissue complications after radiotherapy and polymorphic variations in TGFB1 and XRCC1 genes. *Radiat Res* 2010;173:505–511.
24. Kerns SL, Ostrer H, Stock R, *et al*. Genome-wide association study to identify single nucleotide polymorphisms (SNPS) associated with the development of erectile dysfunction in African-American men after radiotherapy for prostate cancer. *Int J Radiat Oncol Biol Phys* 2010;78:1292–1300.
25. Mangoni M, Bisanzi S, Carozzi F, *et al*. Association between genetic polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT genes and radiosensitivity in breast cancer patients. *Int J Radiat Oncol Biol Phys*. In press.
26. Fogarty GB, Muddle R, Sprung CN, *et al*. Unexpectedly severe acute radiotherapy side effects are associated with single nucleotide polymorphisms of the melanocortin-1 receptor. *Int J Radiat Oncol Biol Phys* 2010;77:1486–1492.
27. Barnett GC, Coles CE, Burnet NG, *et al*. No association between SNPs regulating TGF-beta 1 secretion and late radiotherapy toxicity to the breast: Results from the RAPPER study. *Radiother Oncol* 2010;97:9–14.
28. Hildebrandt MA, Komaki R, Liao Z, *et al*. Genetic variants in inflammation-related genes are associated with radiation-induced toxicity following treatment for non-small cell lung cancer. *PLoS One* 2010;5:e12402.
29. Zhang L, Yang M, Bi N, *et al*. ATM polymorphisms are associated with risk of radiation-induced pneumonitis. *Int J Radiat Biol Oncol Phys* 2010;77:1360–1368.
30. Zhang L, Yang M, Bi N, *et al*. Association of TGF-β1 and XPD polymorphisms with severe acute radiation-induced esophageal toxicity in locally advanced lung cancer patients treated with radiotherapy. *Radiother Oncol* 2010;97:19–25.
31. Yuan XL, Liao ZX, Liu ZS, *et al*. Single nucleotide polymorphism at rs1982073: T869C of the TGF beta 1 gene is

associated with the risk of radiation pneumonitis in patients with non-small-cell lung cancer treated with definitive radiotherapy. *J Clin Oncol* 2009;27:3370–3378.
32. Werbrouck J, De Ruyck K, Duprez F, *et al*. Acute normal tissue reactions in head-and-neck patients treated with IMRT: Influence of dose and association with genetic polymorphisms in DNA DSB repair genes. *Int J Radiat Oncol Biol Phys* 2009; 73:1187–1195.
33. Cesaretti JA, Stock RG, Atencio DP, *et al*. A genetically determined dose-volume histogram predicts for rectal bleeding among patients treated with prostate brachytherapy. *Int J Radiat Oncol Biol Phys* 2007;68:1410–1416.
34. Damaraju S, Murray D, Dufour J, *et al*. Association of DNA repair and steroid metabolism gene polymorphisms with clinical late toxicity in patients treated with conformal radiotherapy for prostate cancer. *Clin Cancer Res* 2006;12:2545–2554.
35. National Cancer Institute Cancer Therapy Evaluation Program. Protocol development. Available at: http://ctep.cancer.gov/reporting/ctc.html. Accessed December 12, 2003.
36. Ovcharenko I, Nobrega MA, Loots GG, *et al*. ECR Browser: A tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res* 2004;32: W280–W286.
37. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B* 1996;58:267–288.
38. Robins JM, Wang N. Inference for imputation estimators. *Biometrika* 2000;87:113–124.
39. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
40. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996;91:473–518.
41. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
42. van Buuren S, Oudshoorn CGM. Multivariate imputation by chained equations. MICE V1.0 User's manual. *TNO report* 2000;PG/VGZ/00.038.
43. Werner-Wasik M, Yorke E, Deasy J, *et al*. Radiation dose-volume effects in the esophagus. *Int J Radiat Oncol Biol Phys* 2010;76:S86–S93.
44. Hardison RC. Conserved noncoding sequences are reliable guides to regulatory elements. *Trends Genet* 2000;16:369–372.
45. Mendelsohn J, Baselga J. Epidermal growth factor receptor targeting in cancer. *Semin Oncol* 2006;33:369–385.
46. McKibbin T, Zhao W, Tagen M, *et al*. Epidermal growth factor receptor polymorphisms and risk for toxicity in paediatric patients treated with gefitinib. *Eur J Cancer* 2010;46: 2045–2051.
47. Giovannetti E, Zucali PA, Peters GJ, *et al*. Association of polymorphisms in AKT1 and EGFR with clinical outcome and toxicity in non-small cell lung cancer patients treated with gefitinib. *Mol Cancer Ther* 2010;9:581–593.
48. Bernabeu C, Blanco FJ, Langa C, *et al*. Involvement of the TGF-beta superfamily signalling pathway in hereditary haemorrhagic telangiectasia. *J Appl Biomed* 2010;8:169–177.
49. Scharpfenecker M, Floot B, Russell NS, *et al*. Endoglin haploinsufficiency reduces radiation-induced fibrosis and telangiectasia formation in mouse kidneys. *Radiother Oncol* 2009; 92:484–491.
50. Kruse JJCM, Floot BGJ, te Poele JAM, *et al*. Radiation-induced activation of TGF-beta signaling pathways in relation to vascular damage in mouse kidneys. *Radiat Res* 2009;171: 188–197.
51. Wang YY, Zhang P, Liu YF, *et al*. TRAF-mediated regulation of immune and inflammatory responses. *Sci China Ser C* 2010;53:159–168.
52. Baluna RG, Eng TY, Thomas CR. Adhesion molecules in radiotherapy. *Radiat Res* 2006;166:819–831.