Clinical Investigation

A Validated Prediction Model for Overall Survival From Stage III Non-Small Cell Lung Cancer: Toward Survival Prediction for Individual Patients

CrossMark

Cary Oberije, PhD,* Dirk De Ruysscher, MD, PhD,*,†
Ruud Houben, MSc,* Michel van de Heuvel, MD, PhD,‡
Wilma Uyterlinde, MSc, PhD,‡ Joseph O. Deasy, PhD,§
Jose Belderbos, MD, PhD,∥ Anne-Marie C. Dingemans, MD, PhD,¶
Andreas Rimner, MD, PhD,§ Shaun Din, MD,§ and
Philippe Lambin, MD, PhD*

*Radiation Oncology, Research Institute GROW of Oncology, Maastricht University Medical Center, Maastricht, The Netherlands; †Universitaire Ziekenhuizen Leuven, KU Leuven, Belgium; ‡Department of Thoracic Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands; §Memorial Sloan Kettering Cancer Center, New York; ∥Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands; and ¶Department of Pulmonology, University Hospital Maastricht, Research Institute GROW of Oncology, Maastricht, The Netherlands

Summary

Although patients with stage III non-small cell lung cancer are homogeneous according to the TNM staging system, they form a heterogeneous group. General guidelines should be replaced by personalized

Purpose: Although patients with stage III non-small cell lung cancer (NSCLC) are homogeneous according to the TNM staging system, they form a heterogeneous group, which is reflected in the survival outcome. The increasing amount of information for an individual patient and the growing number of treatment options facilitate personalized treatment, but they also complicate treatment decision making. Decision support systems (DSS), which provide individualized prognostic information, can overcome this but are currently lacking. A DSS for stage III NSCLC requires the development and integration of multiple models. The current study takes the first step in this process by developing and validating a model that can provide physicians with a survival probability for an individual NSCLC patient.

treatment, but this requires individualized prognostic information. Demographic, tumor, and treatment data from 548 patients were used to develop a model for predicting the survival of individual patients. The model was validated in 2 external data sets. The results can be used for survival prediction, and the baseline model can be further improved in the future.

Methods and Materials: Data from 548 patients with stage III NSCLC were available to enable the development of a prediction model, using stratified Cox regression. Variables were selected by using a bootstrap procedure. Performance of the model was expressed as the c statistic, assessed internally and on 2 external data sets (n = 174 and n = 130).

Results: The final multivariate model, stratified for treatment, consisted of age, gender, World Health Organization performance status, overall treatment time, equivalent radiation dose, number of positive lymph node stations, and gross tumor volume. The bootstrapped c statistic was 0.62. The model could identify risk groups in external data sets. Nomograms were constructed to predict an individual patient's survival probability (www.predictcancer.org). The data set can be downloaded at https://www.cancerdata.org/10.1016/j.ijrobp.2015.02.048.

Conclusions: The prediction model for overall survival of patients with stage III NSCLC highlights the importance of combining patient, clinical, and treatment variables. Nomograms were developed and validated. This tool could be used as a first building block for a decision support system. © 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

In Europe, lung cancer is by far the most common cause of cancer death in men and the third cause of cancer deaths in women (1), and in the United States, lung cancer death holds the first position for both sexes (2). In 2012, more than 400,000 new cases were diagnosed in Europe. Approximately 30% of patients with non-small cell lung cancer (NSCLC) receive a diagnosis of stage III disease.

The heterogeneity in this patient group makes it difficult to choose the optimal treatment for an individual patient (3). Moreover, this heterogeneity is becoming more prominent as new imaging modalities, genomics, and proteomics approaches are being used to describe tumors and patients. In addition, the number of treatment options is rising and includes individualized chemotherapy, targeted agents, new radiation therapy schemes and techniques, proton therapy, surgery, or a combination of these options. A decision support system (DSS) could offer assistance for treatment decision making but is currently lacking. This system should incorporate multiple models to predict several relevant outcomes for different treatment options (4) (Fig. E1, available online at www.redjournal.org). A model that consists of basic clinical variables and predicts survival outcome for individual patients could serve as a first building block for this DSS. In addition, more accurate prediction of survival would allow identification of patients with comparable prognoses and could be useful for risk stratification in clinical trials. Also, doctors and patients would have better information about the prognosis and could take this into account in a shared decision making process.

During recent decades, numerous studies have investigated prognostic and predictive factors for lung cancer survival. By contrast, studies especially focusing on stage III NSCLC are relatively scarce (5).

The aim of this study was to develop and validate a prediction model for survival of stage III NSCLC patients, treated with (chemo) radiation therapy, taking into account all available and established prognostic factors.

Methods and Materials

Patient population

Between March 2002 and August 2011, data were collected prospectively for several patient cohorts (NCT00181545 clinicaltrials.gov, NCT00181506 clinicaltrials.gov, NCT00572325 clinicaltrials.gov, NCT00573040 clinicaltrials.gov, NCT01166204 clinicaltrials.gov, NCT01084785 clinicaltrials.gov, NCT01936571 clinicaltrials.gov), ensuring standardization and high quality of data. All patients were treated with radiation therapy with curative intent at the MAASTRO Clinic. For the current analysis, all inoperable patients with stage III NSCLC (according to the 6th edition of the TNM staging system) were selected. Patients were excluded from the study if (*1*) they had received a diagnosis of another primary tumor less than 5 years ago; (*2*) positron emission tomography (PET) was not used for staging; and (*3*) they had malignant pleural effusion. In addition, 2 patients were lost to follow-up, 2 patients refused treatment, and 4 patients died during the course of radiation therapy. The statistical analysis is based on 548 patients. The primary gross tumor volume ($GTV_{primary}$) and the nodal gross tumor volume (GTV_{nodal}) were delineated manually, using information from PET and computed tomography (CT). For patients treated with sequential chemotherapy, these volumes were calculated using postchemotherapy imaging information. The number of positive lymph node stations was assessed by the nuclear medicine specialist using either an

integrated ^{18}F-fluorodeoxyglucose (FDG)-PET-CT scan or a CT scan combined with FDG-PET scan. T stage and N stage were assessed using pretreatment CT, PET, and mediastinoscopy when applicable. For patients treated with sequential chemotherapy, stage and the number of positive lymph node stations were assessed using prechemotherapy imaging information.

Radiation therapy development cohort

All patients were treated at the MAASTRO Clinic, using CT-based radiation therapy planning. No elective nodal irradiation was performed, and irradiation was delivered 5 days per week. Patients were treated with 3-dimensional conventional radiation therapy (3D-CRT) (2002 to January 2010) or intensity modulated radiation therapy (IMRT) (February 2010 to 2011).

Four different radiation treatment regimens were applied.

1. Twenty-five patients were included in a phase 1 dose escalation study (6). They received a total radiation dose ranging from 61.2 to 68.4 Gy, delivered twice daily in fractions of 1.8 Gy.

2. The second group consisted of 135 patients who were treated according to the standard protocol, used until August 2005. They received 60 Gy in daily fractions of 2 Gy.

3. One hundred eighty-seven patients were treated according to the standard protocol that was introduced in August 2005 (7). The radiation dose ranged from 54.0 to 79.2 Gy, delivered in fractions of 1.8 Gy twice daily, depending on the mean lung dose or the spinal cord dose constraint.

4. Since August 2008, concurrent chemoradiation was delivered to patients judged to be physically fit to undergo this treatment (8). A total of 201 patients received a total dose of 45 Gy, delivered in 1.5 Gy-fractions twice daily. Subsequently, a dose ranging from 4.0 to 24.0 Gy was delivered in 2.0 Gy fractions once daily, depending on the mean lung dose and the spinal cord dose constraint.

The equivalent dose in 2-Gy fractions (EQD$_2$) was used as a measure for the intensity of chest radiation therapy delivered to the tumor (9). Adjustment for dose per fraction was made as follows:

$$EQD_2 = D * \left(\frac{d + \frac{\alpha}{\beta}}{2 + \frac{\alpha}{\beta}} \right)$$

where D = the total radiation dose, d = dose per fraction, α/β = 10 Gy.

Chemotherapy

Two hundred eighty-two patients, treated according to regimens 1 through 3, received chemotherapy before radiation therapy started (Table 1). The regimen consisted of carboplatin on day 1 and gemcitabine on days 1 and 8. The majority received 3 cycles (range, 1-6). Patients treated according to regimen 4 received concurrent chemoradiation, which consisted of 1 to 2 cycles of carboplatin and gemcitabine, followed by concurrent cisplatin-vinorelbine or concurrent cisplatin-etoposide every 3 weeks with radiation therapy. The regimen depended on the referring hospital. Dose reduction was applied according to guidelines and in case of renal failure cisplatin was substituted by carboplatin.

Survival endpoint

Overall survival was defined as the duration between the start of radiation therapy and the date of death. Survival status for the development cohort was evaluated in December 2013 using a population registration system. Data were considered right-censored if patients were alive at the time point of evaluation.

Statistical analysis

The Kaplan-Meier method was used for univariate survival analysis. To build a multivariate prediction model, Cox regression was applied. It was assumed that the treatment choice partially depended on the characteristics of the patient cohorts. To correct for this so-called "confounding by indication," the analysis was stratified by treatment cohort: (1) radiation therapy only; (2) treated with sequential chemoradiation according to standard protocol; (3) treated with concomitant chemoradiation according to standard protocol; and (4) sequential chemoradiation for patients who were not eligible for concomitant chemoradiation (Fig. E2, available online at www.redjournal.org). Continuous variables were modeled nonlinearly by applying restricted cubic splines (10, 11). If possible, the nonlinear terms were replaced by simpler transformations afterward. Variable selection was performed with a bootstrap procedure. First, it was determined how many variables should be included in the model by performing backwards variable selection for 100 bootstrap samples. Subsequently, backwards selection on the original data set was applied to identify which variables should be retained in the model. The final model was compared with 2 models that included variables based on the TNM staging system. The performance of the models was assessed in terms of model fit and discrimination. Discrimination (ie ability to classify patients correctly) was assessed using the c statistic. Its interpretation is comparable to the AUC: the maximum value is 1.0 and indicates a perfect prediction model, and a value of 0.5 indicates that patients are

Table 1 Patient characteristics

Variable	Development cohort MAASTRO Clinic (n=548)	Validation cohorts NKI (n=174)	P*	MSKCC (n=130)	P*
Mean age (y)	66 (SD 10)	63 (SD 10)	<.001	67 (SD 11)	.510
Gender			.144		<.001
Male	379 (69.2%)	110 (63.2%)		63 (48.5%)	
Female	169 (30.8%)	64 (36.8%)		67 (51.5%)	
WHO-PS			<.001		.013
0	192 (35.0%)	-		48 (36.9%)	
1	287 (52.4%)	139 (79.9%)		78 (60.0%)	
≥2	63 (11.6%)	35 (20.1%)		4 (3.1%)	
Missing	6 (1.0%)	-		-	
Mean FEV_1 (%)	76 (range 21-139)	78 (range 37 - 133)	.292		
Missing	71 (13.0%)	34 (19.5%)			
BMI	24.9 (SD 4.3)	NA	-	NA	
Missing	179 (32.7%)				
Nicotine use		NA	-		<.001
No/former smoker	304 (55.5%)			100 (76.9%)	
Current smoker	202 (36.9%)			30 (23.1%)	
Missing	42 (7.7%)			-	
Clinical T stage			.053		<.001
T1	74 (13.5%)	28 (16.1%)		27 (20.8%)	
T2	172 (31.4%)	67 (38.5%)		42 (32.3%)	
T3	60 (10.9%)	27 (15.5%)		27 (20.8%)	
T4	216 (39.4%)	52 (29.9%)		31 (23.8%)	
Missing	26 (4.7%)	-		3 (2.3%)	
Clinical N stage			<.001		
N0	95 (17.3%)	13 (7.5%)			
N1	15 (2.7%)	7 (4.0%)			
N2	267 (48.7%)	123 (70.7%)			
N3	167 (30.5%)	31 (17.8%)			
Missing	4 (0.7%)				
Clinical overall stage			<.001	NA	
IIIA	199 (36.3%)	115 (66.1%)			
IIIB	349 (66.1%)	58 (33.3%)			
Missing		1 (0.6%)			
Histology			<.001		
Adenocarcinoma	81 (14.8%)	35 (20.1%)			
SCC	164 (29.9%)	54 (31.0%)			
Large cell carcinoma	190 (34.7%)	78 (44.8%)			
Other	93 (17.0%)	7 (4.0%)			
Unknown	20 (3.7%)	-			
Median GTV (mL) (range)	51 (0-725)	99 (0-1822)	<.001[†]	54 (0.3-1057)	.360[†]
Missing	41 (7.5%)	5 (2.9%)		-	
PLNS			.603		<.001
0	104 (19.0%)	36 (20.7%)		6 (4.6%)	
1	107 (19.5%)	36 (20.7%)		14 (10.8%)	
2	113 (20.6%)	32 (18.4%)		39 (30.0%)	
3	69 (12.6%)	31 (17.8%)		36 (27.7%)	
≥4	125 (22.9%)	39 (22.4%)		35 (26.9%)	
Missing	30 (5.5%)	-			
Chemotherapy			-		.009
No	66 (12.0%)	-		12 (9.2%)	
Sequential	280 (51.1%)	-		51 (39.2%)	
Concurrent	202 (36.9%)	174 (100%)		67 (51.5%)	
Mean OTT (d)	32 (SD 8)	31 (SD 2)	.763	43 (SD 8)	<.001
Mean EQD_2 (Gray)	60.8 (SD 7.1)	70.1 (SD 0)	-	60.1 (SD 8.8)	.313

Abbreviations: BMI = body mass index; EQD_2 = equivalent radiation dose at 2 Gy; FEV_1 = forced expiratory volume in 1 second; GTV = gross tumor volume; MSKCC = Memorial Sloan Kettering Cancer Center; NKI = Netherlands Cancer Institute; OTT = overall treatment time radiation therapy; PLNS = number of positive lymph node stations; SCC = squamous cell carcinoma; SD = standard deviation; WHO-PS = World Health Organization performance status.

 * P values assessed using *t* test (continuous variables) or χ^2 test (categorical variables).

 [†] P values assessed using Mann-Whitney test (nonparametric distribution).

correctly classified in 50% of the cases (eg as good as chance). To correct for a too optimistic estimation of the c statistic, bootstrapping was applied (11). Akaike's information criterion (AIC) is a trade-off between model fit and number of included variables and can be used to compare nonnested models. The preferred model is the one with the lowest AIC value. The difference can be interpreted as follows: a decrease by 4 to 7 indicates weak support; a decrease by 10 indicates strong support for a model. Differences of 0 to 3 indicate that performance is comparable (12, 13). To assess the model performance on the external data sets, the cohort was split into 3 subgroups, based on 25th and 75th percentile of the risk score (14). For these subgroups a calibration plot, in which the predicted probability is compared with the observed outcome, and Kaplan-Meier curves were made. Nomograms were made for practical use. The analysis was performed with SPSS, version 19.0 (SPSS Inc, Chicago, IL) and R, version 2.14.0, using the RMS library (Harrell).

External validation cohorts

The validation cohorts consisted of 174 and 130 patients with stage III NSCLC treated with sequential or concurrent chemoradiation, or radiation therapy alone, at the Netherlands Cancer Institute (NKI) and Memorial Sloan Kettering Cancer Center (MSKCC), respectively. Staging was performed using FDG-PET information. All NKI patients received definitive concurrent chemoradiation with IMRT (Table 1) (15). No elective irradiation was performed. Treatment consisted of 24 fractions of 2.75 Gy and daily cisplatin. The MSKCC cohort was treated with conventional 3D radiation therapy or IMRT, and chemotherapy consisted of a cisplatin doublet or, less commonly, carbo/taxol every 3 weeks (16, 17).

Ethics

This study was conducted according to national laws and guidelines and approved by the local review boards.

Results

Development cohort

Table 1 depicts the characteristics of the development cohort and the validation cohort. Of all patients, 456 (83.2%) had died at the time of the analysis. The median follow-up time was 5.5 years (range, 1.3-10.0 years). The median survival for the whole group was 16 months (95% confidence interval [CI]: 14-19 months), with a 2-year survival of 39% (95% CI: 32%-46%) for stage IIIA disease and 36% (95% CI: 31%-41%) for stage IIIB disease (log-rank $P = .84$). The associations between EQD_2 and OTT or GTV are shown in Figure E8 (available online at www.redjournal.org).

External validation cohort

There were several statistically significant differences between the MAASTRO Clinic cohort and the validation cohorts (Table 1). The mean GTV of patients treated at the NKI was larger, and the radiation therapy dose was significantly higher. In addition, the NKI cohort was more homogeneous with respect to World Health Organization (WHO) performance status, chemotherapy, OTT, and EQD_2. The median follow-up time was 28 months (range, 2-49 months). The 2-year overall survival was 51% (95% CI: 43%-59%). The MSKCC cohort consisted of a higher percentage of women, fewer smokers, and fewer T4 tumors. There were only a few patients with N0. Compared with the MAASTRO Clinic cohort, more patients received concurrent chemoradiation, and the OTT was longer. The median follow-up time was 62 months (range, 6-97 months). The 2-year overall survival was 44% (95% CI: 36%-54%).

Multivariate model

The variables available for model building were treatment cohort (stratification factor), age, gender, WHO performance status, body mass index, forced expiratory volume (in 1 second), clinical overall stage, T stage, N stage, number of positive lymph node stations (PLNS), GTV, OTT, and EQD_2. After the variable selection procedure, the final model consisted of gender, WHO performance status, T stage, GTV, PLNS, OTT, and EQD_2 (Table 2). Age was included in the model although it was statistically not significant. The association of GTV, OTT, and age with survival was modeled nonlinearly. The hazard of death did not further increase for GTVs larger than approximately 150 mL, nor did the hazard further decrease for treatment times shorter than approximately 28 days or age <70 (Fig. E3, available online at www.redjournal.org). These nonlinear relationships could be approximated very well by a log transformation for GTV and a piecewise linear function, consisting of 2 straight-line sections, for OTT and age (Fig. 1). These simplifications did not result in a decreased performance. The c statistic of the final model was 0.65 (0.62 after correction for optimism). Comparison with 2 models that included variables based on the TNM staging system showed that model fit and predictive performance were better for the multivariate model (Table 3). Application of the model on the data sets of the NKI and MSKCC resulted in c statistics of 0.58 and 0.60, respectively. Although the differences between observed and predicted survival probabilities were small for the NKI cohort, there were considerable disparities for the MSKCC cohort (Fig. E4, available online at www.redjournal.org). The model underestimated overall survival for the MSKCC patients. To create Kaplan-Meier curves, the validation cohorts were split into 3 subgroups based on the 25th and 75th percentile of the predicted probability (14). For the NKI cohort, a high-risk group could be identified, but the

Table 2   Hazard ratios for overall survival*

Variable	Values	Coef	SE	HR	95% CI	P
Sex	Male			ref		.011
	Female	−0.28	0.11	0.76	0.61-0.94	
Age (y)	≤70			ref		
		0.024	0.015	1.02	0.99-1.05	.1402
WHO-PS	0			ref		.0008
	1	0.31	0.11	1.37	1.11-1.68	
	≥2	0.52	0.16	1.68	1.22-2.31	
PLNS	0			ref		<.0001
	1	0.27	0.16	1.31	0.95-1.80	
	2	0.49	0.16	1.63	1.19-2.24	
	3	0.67	0.18	1.96	1.38-2.79	
	≥4	0.79	0.18	2.21	1.55-3.14	
T stage	T0/T1			ref		.3135
	T2	0.10	0,16	1,11	0.81-1.52	
	T3	0.32	0,21	1.8	0.92-2.07	
	T4	0.06	0.17	1.06	0.76-1.50	
LN GTV (mL)		0.15	0.01	1.16	1.14-1.18	.0008
OTT (days)	≤28			ref		
		0.040	0.009	1.04	1.02-1.06	<.0001
EQD$_2$		−0.015	0.008	0.99	0.97-1.00	.0506

Abbreviations: CI = confidence interval; Coef = regression coefficient; EQD$_2$ = equivalent radiation dose at 2 Gy; GTV = gross tumor volume; LN = natural logarithm; OTT = overall treatment time radiation therapy; PLNS = number of positive lymph node stations; SE = standard error; WHO-PS = World Health Organization performance status.

* Model is stratified by treatment.

difference between the other groups was small (log-rank $P = .0006$) (Fig. 2A), and identification of a low-risk group was possible for the MSKCC cohort (log-rank $P = .059$) (Fig. 2B). Nomograms were developed for prediction of 24 months (Fig. 3) and 36 months overall survival (Fig. E5, available online at www.redjournal.org). The model will be freely available on www.predictcancer.org. In addition to the point estimate, 95% confidence intervals are calculated on the website. The data set is publicly available on https://www.cancerdata.org/10.1016/j.ijrobp.2015.02.048.

Discussion

To our knowledge, this is the first prediction model for overall survival of patients with stage III NSCLC treated with (chemo) radiation therapy. The performance of the model in the development cohort and in the validation data sets was moderate, but evidently better than the performance of the TNM-based models. We showed that accurate risk stratification for these patients should be based on multiple factors, including WHO performance status, gender, PLNS, GTV, clinical T stage, chemotherapy, OTT, and EQD$_2$.

In agreement with the literature, where performance status, regardless of resectability of the tumor, is consistently identified as an important predictor for survival (18, 19), we found that worse performance status was associated with shorter survival.

Although some studies found no influence of gender (20), there are also indications that female gender is a

favorable factor for survival (21, 22). In our study, female gender was prognostic for better survival.

Several prospective studies have suggested that higher radiation doses lead to improved local control and higher survival rates (7, 8, 23). However, the randomized RTOG 0617 trial, comparing 60 Gy with 74 Gy, showed a significantly lower survival in the high-dose arm (24). This unexpected finding could be explained by any of several factors, including the prolonged OTT and increased heart toxicity in the high-dose group (25). Previously published studies have reported a decrease in tumor control probability of 1.6% per day after a 6-week duration of radiation therapy (26) and a 2.0% increase in the risk of death for each day of prolongation in therapy (27). This is due to accelerated repopulation, which occurs if OTT is longer than approximately 35 to 42 days. In our study, patients were treated with a wide range of doses and accelerated repopulation was avoided by restricting the OTT to a maximum of approximately 40 days (7, 8). Obviously, this heterogeneity made it possible to find and model the dose-response relationship. The difference in 2-year survival (44% vs 51%) between the patients treated with concurrent chemotherapy from the MAASTRO Clinic cohort and those from the NKI cohort could possibly be explained by the radiation dose scheme (9 Gy lower in the MAASTRO Clinic cohort), indicating a beneficial effect of dose intensification for these patients. Previously published studies have reported a slightly increased survival for intermediately escalated dose (28, 29). The survival difference between MAASTRO Clinic and MSKCC might be attributed to a higher percentage of women and a lower percentage of smokers. Also, differences in (introduction of) diagnostic
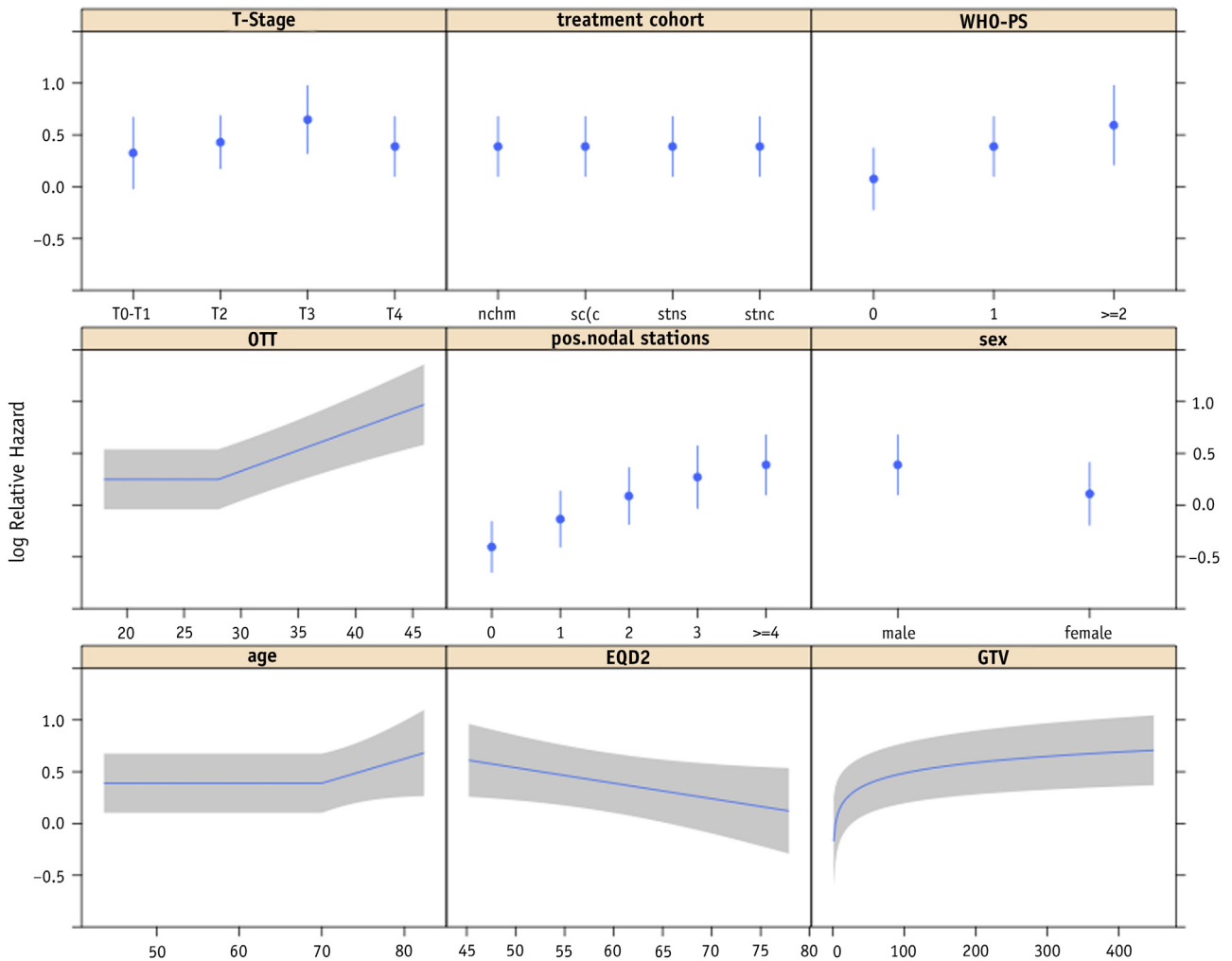
Fig. 1. Predictors in the final multivariable model, using transformed variables to simplify the model. No regression coefficients are estimated for the treatment cohorts as overall survival is estimated by the Kaplan-Meier method for the strata (treatment cohorts). EQD_2 = equivalent dose in 2-Gy fractions; GTV = gross tumor volume; OTT = overall treatment time; WHO-PS = World Health Organization performance status.

procedures such as magnetic resonance imaging for the detection of brain metastases, or chemotherapy regimens for both primary and second-line treatment, could have influenced the case mix and thus explain our results. The fact that the accuracy of the model was higher for the NKI cohort from the Netherlands than for the MSKCC cohort from the United

States indicates that these patient populations are not completely comparable, but more research is needed to shed light on the possible explanations (Fig. E4, available online at www.redjournal.org).

Several groups have concluded that GTV is a highly significant factor for the prediction of survival (30, 31). The nonlinear relationship has been reported previously by the MAASTRO Clinic (32) but also by others (30, 33). In addition, Ball et al (33) reported a time-dependent hazard ratio for the GTV, but this finding was not replicated in our data (32). The GTV was the most important prognostic factor in our analysis.

The importance of a short OTT for survival outcome has been reported in several studies (27, 34, 35). From a radiobiological point of view it can be assumed that treatment times shorter than 28 days do not further improve the survival of NSCLC patients (9). Our clinical data support this assumption.

Previously it has already been shown that PLNS on FDG-PET scan is an important risk factor for

Table 3 Comparison of performance and model fit

Model	Performance*	95% CI	$P^{\dagger}$	AIC
MV model with rcs	0.62	0.61-0.66		3945
Simplified MV model	0.62	0.61-0.66		3945
Stage IIIA vs stage IIIB	0.54	0.52-0.58	<.001	4001
T stage + N stage	0.57	0.56-0.61	<.001	3991

Abbreviations: AIC = Aikaike's information criterion; CI = confidence interval; MV = multivariable; rcs = restricted cubic splines.
 * Performance assessed by bootstrap procedure.
 † Comparison between simplified MV model and models based on TNM stage.

A

B
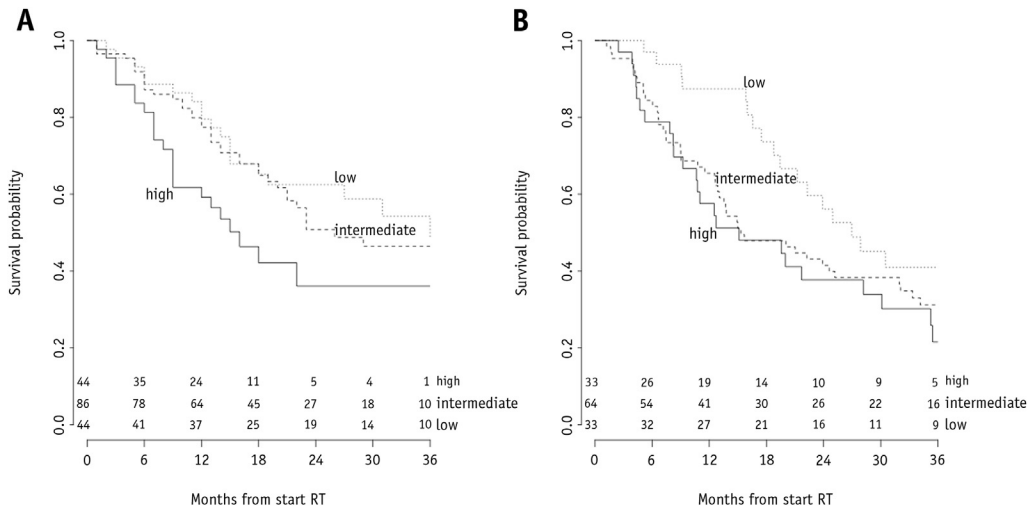
Fig. 2. Kaplan-Meier curves for overall survival of risk groups, based on the predicted probability for the external validation cohorts from (A) Y (n=174) and (B) Z (n=130). RT = radiation therapy.

non-surgically treated patients (32, 36). The current study confirms these results. Although a recently updated meta-analysis suggests that high standardized uptake value (SUV) measured by FDG-PET is a poor prognostic factor for survival in NSCLC patients, its role is less clear for stage III patients (5). In addition, comparison of SUVs

across institutes and countries is difficult because of the lack of standardization. Therefore, using PLNS assessed by FDG-PET instead of SUV-based measures enhances the applicability of our model.

It can be concluded that the performance of our model, c statistic of 0.62, is moderate, but evidently better than the
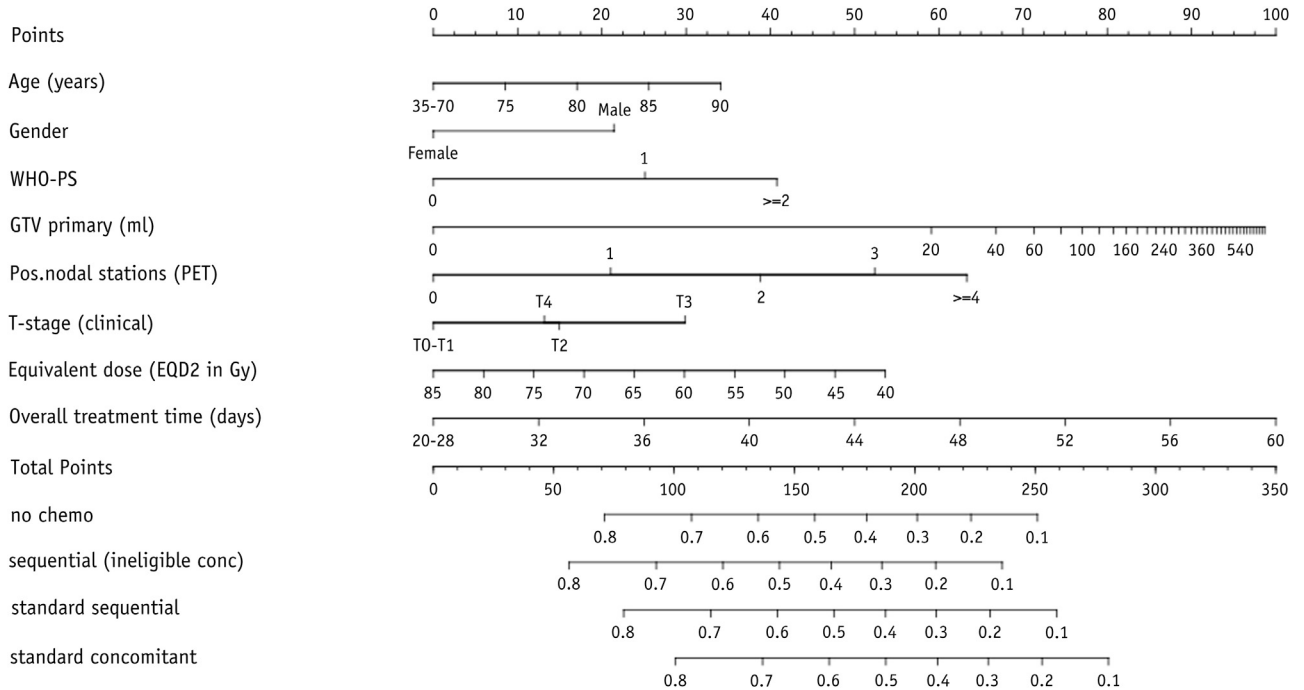
Fig. 3. Nomogram for prediction of 24-month overall survival. The outcome is a point estimate; 95% confidence intervals can be obtained from the website www.predictcancer.org. *Instructions for physician*: Locate the patient's age on the age axis. Draw a line straight upward to the points axis to determine how many points a patients receives for age. Repeat this process for the other axes, each time drawing straight upward to the points axis. Sum the points achieved for each predictor, and locate this sum on the "Total points" axis. Draw a line straight down to assess the survival probability for this patient. EQD_2 = equivalent dose in 2-Gy fractions; GTV = gross tumor volume; PET = positron emission tomography; WHO-PS = World Health Organization performance status.

predictive performance of the models that included only overall stage or T stage and N stage (c statistic 0.50 and 0.55, respectively). The performance of these models increased if treatment cohort was entered as a stratification factor (0.54 and 0.57, respectively).

Given that our model was based on patients treated with 3D-CRT or IMRT, predictions for patients treated with other techniques or treatment schemes should be interpreted carefully (for an example, See Text E7, available online at www.redjournal.org). Moreover, the majority of patients were treated according to a nonstandard dose escalation protocol, which severely limits the applicability of the nomogram. Therefore, the nomogram should not be used to adapt treatment of individual patients. In addition, only a point estimate is provided to indicate the predicted probability, which can be misleading. It is possible to obtain 95% confidence intervals from the website (www. predictcancer.org).

The cutoff points to define the risk groups were predefined to avoid over optimistic results. However, clinically more relevant cutoff points could be chosen to select patients for clinical trials or for specific treatment modalities such as proton therapy (Fig. E1, available online at www.redjournal.org).

Although predictive factors define the effect of treatment on the outcome, prognostic factors indicate the effects of patient or tumor characteristics. Inasmuch as our model contains treatment characteristics that can be adapted, we decided to use the term "prediction model," although some might opt for the term "prognostic."

We validated our model in 2 external data sets, resulting in c statistics of 0.59 and 0.60. The decrease was thus limited and can probably be attributed to differences in case mix. Moreover, the Kaplan-Meier curves show that it was possible to identify risk groups in the validation cohorts (Fig. 2).

The treatment of the patient cohort included in this study was very heterogeneous, and the choice of treatment regimen was partially guided by patients' general health and physical fitness, leading to "confounding by indication." To adjust the survival curve for factors that cannot be captured in the prognostic variables, a stratified analysis was performed. We were able, despite the heterogeneity of the population, to find several statistically significant and clinically relevant patient and treatment characteristics that were present in all strata.

Other factors with additional prognostic value could further improve the model. Although some biomarkers prognostic for survival have already been identified, these results still have to be confirmed and validated on a larger scale. In addition, genomics and proteomics analysis hold great promise but still have their limitations in clinical application (5, 37). Finally, radiomics, sophisticated analysis of imaging information, might provide new prognostic or predictive factors (38, 39).

Although most studies focus on a specific group of variables (eg biomarkers, genes, imaging, or clinical

characteristics) in our opinion integration of information from multiple sources is the way forward toward more accurate prediction models. For that purpose, large databases, new modeling methods, and a culture of data sharing are needed because it will be impossible to collect all this information from the same group of patients or within 1 trial (40, 41).

The current model can be considered a first building block of a DSS (4). Ultimately, this DSS should be tested prospectively in a clinical trial. In our opinion, it is only a matter of time until DSSs are introduced into clinical practice to offer assistance for treatment decision making by quantifying the risks and the benefits of a specific treatment.

Although new factors are needed to improve the predictive capability of the survival model, and patient cohorts treated with standard radiation dose should be included to increase the applicability, our results are encouraging.

It is concluded that in this study, we developed and validated a survival model for patients with stage III NSCLC treated with (chemo) radiation. The model, which estimates 2-year and 3-year survival of individual patients, outperforms TNM-based models and could therefore provide clinicians and patients with more specific prognostic information.

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *Eur J Cancer* 2013;49:1374-1403.
2. Siegel R, Ma J, Zou Z, et al. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9-29.
3. Vansteenkiste J, De Ruysscher D, Eberhardt WE, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24(Suppl 6):vi89-vi98.
4. Lambin P, van Stiphout RG, Starmans MH, et al. Predicting outcomes in radiation oncology—multifactorial decision support systems. *Nat Rev Clin Oncol* 2013;10:27-40.
5. Berghmans T, Paesmans M, Sculier JP. Prognostic factors in stage III non-small cell lung cancer: A review of conventional, metabolic and new biological variables. *Ther Adv Med Oncol* 2011;3:127-138.
6. De Ruysscher D, Wanders R, van Haren E, et al. HI-CHART: A phase I/II study on the feasibility of high-dose continuous hyperfractionated accelerated radiotherapy in patients with inoperable non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2008;71:132-138.
7. van Baardwijk A, Wanders S, Boersma L, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I to III non-small-cell lung cancer. *J Clin Oncol* 2010;28:1380-1386.
8. van Baardwijk A, Reymen B, Wanders S, et al. Mature results of a phase II trial on individualised accelerated radiotherapy based on normal tissue constraints in concurrent chemo-radiation for stage III non-small cell lung cancer. *Eur J Cancer* 2012;48:2339-2346.
9. Fowler JF. Biological factors influencing optimum fractionation in radiation therapy. *Acta Oncol* 2001;40:712-717.
10. Steyerberg EW. Clinical Prediction Models. New York, NY: Springer New York; 2009.
11. Harrell FE Jr. Regression Modeling Strategies. New York: Springer-Verlag; 2001.
12. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;19:716-723.

13. Burnham KP, Anderson DR. Model Selection and Multimodel Inference. New York: Springer Science & Business Media; 2002.

14. Cox DR. Note on grouping. *J Am Stat Assoc* 1957;52:543-547.

15. Uyterlinde W, Belderbos J, Baas C, et al. Prediction of acute toxicity grade ≥3 in patients with locally advanced non-small cell lung cancer receiving intensity modulated radiotherapy and concurrent low-dose cisplatin. *Clin Lung Cancer* 2013;14:541-548.

16. Rosenzweig KE, Sura S, Jackson A, et al. Involved-field radiation therapy for inoperable non smallcell lung cancer. *J Clin Oncol* 2007; 25:5557-5561.

17. Sura S, Gupta V, Yorke E, et al. Intensity modulated radiation therapy (IMRT) for inoperable non-small cell lung cancer: The Memorial Sloan-Kettering Cancer Center (MSKCC) experience. *Radiother Oncol* 2008;87:17-23.

18. Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: A decade of progress. *Chest* 2002;122: 1037-1057.

19. Solan MJ, Werner-Wasik M. Prognostic factors in non-small cell lung cancer. *Semin Surg Oncol* 2003;21:64-73.

20. Werner-Wasik M, Scott C, Cox JD, et al. Recursive partitioning analysis of 1999 Radiation Therapy Oncology Group (RTOG) patients with locally-advanced non-small-cell lung cancer (LA-NSCLC): Identification of five groups with different survival. *Int J Radiat Oncol Biol Phys* 2000;48:1475-1482.

21. Pfister DG, Johnson DH, Azzoli CG, et al. American Society of Clinical Oncology treatment of unresectable non-small cell lung cancer guideline: Update 2003. *J Clin Oncol* 2004;22:330-353.

22. Blanchon F, Grivaux M, Asselain B, et al. 4-year mortality in patients with non-small cell lung cancer: Development and validation of a prognostic index. *Lancet Oncol* 2006;7:829-836.

23. Rengan R, Rosenzweig KE, Venkatraman E, et al. Improved local control with higher doses of radiation in large-volume stage III non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;60: 741-747.

24. Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): A randomised, two-by-two factorial phase 3 study. *Lancet Oncol* 2015;16:187-199.

25. Kong FM, Zhao J, Wang J, et al. Radiation dose effect in locally advanced non-small cell lung cancer. *J Thorac Dis* 2014;6:336-347.

26. Cox JD, Pajak TF, Asbell S, et al. Interruptions of high-dose radiation therapy decrease long-term survival of favorable patients with unresectable non-small cell carcinoma of the lung: Analysis of 1244 cases from 3 Radiation Therapy Oncology Group (RTOG) trials. *Int J Radiat Oncol Biol Phys* 1993;27:493-498.

27. Machtay M, Hsu C, Komaki R, et al. Effect of overall treatment time on outcomes after concurrent chemoradiation for locally advanced non-small cell lung carcinoma: Analysis of the Radiation Therapy Oncology Group (RTOG) experience. *Int J Radiat Oncol Biol Phys* 2005;63:667-671.

28. Rodrigues G, Oberije C, Senan S, et al. Is intermediate radiation dose escalation with concurrent chemotherapy for stage III non−small cell lung cancer beneficial? A multi-institutional propensity score matched analysis. *Int J Radiat Oncol Biol Phys* 2015;91:133-139.

29. Koshy M, Malik R, Sher DJ, et al. The effect of radiotherapy dose on survival in stage III non-small cell lung cancer patients undergoing definitive chemoradiotherapy. *Clin Lung Cancer* 2014;15:365-371.

30. Werner-Wasik M, Swann RS, Bradley J, et al. Increasing tumor volume is predictive of poor overall and progression-free survival: Secondary analysis of the Radiation Therapy Oncology Group 93-11 phase I-II radiation dose-escalation study in patients with inoperable non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2008;70: 385-390.

31. Basaki K, Abe Y, Aoki M, et al. Prognostic factors for survival in stage III non-small cell lung cancer treated with definitive radiation therapy: Impact of tumor volume. *Int J Radiat Oncol Biol Phys* 2006;64: 449-454.

32. Dehing-Oberije C, Yu S, De Ruysscher D, et al. Development and external validation of prognostic model for 2-year survival of non-−small-cell lung cancer patients treated with chemoradiotherapy. *Int J Radiat Oncol Biol Phys* 2008;74:355-362.

33. Ball DL, Fisher RJ, Burmeister BH, et al. The complex relationship between lung tumor volume and survival in patients with non-small cell lung cancer treated by definitive radiotherapy: A prospective, observational prognostic factor study of the Trans-Tasman Radiation Oncology Group (TROG 99.05). *Radiother Oncol* 2013;106:305-311.

34. Baumann M, Herrmann T, Koch R, et al. Final results of the randomized phase III CHARTWEL-trial (ARO 97-1) comparing hyperfractionated-accelerated versus conventionally fractionated radiotherapy in non-small cell lung cancer (NSCLC). *Radiother Oncol* 2011;100:76-85.

35. Uitterhoeve AL, Belderbos JS, Koolen MG, et al. Toxicity of high-dose radiotherapy combined with daily cisplatin in non-small cell lung cancer: Results of the EORTC 08912 phase I/II study. European Organization for Research and Treatment of Cancer. *Eur J Cancer* 2000;36:592-600.

36. Dehing-Oberije C, De Ruysscher D, van der Weide H, et al. Tumor volume combined with number of positive lymph node stations is a more important prognostic factor than TNM stage for survival of non-small-cell lung cancer patients treated with (chemo)radiotherapy. *Int J Radiat Oncol Biol Phys* 2008;70:1039-1044.

37. Dehing-Oberije C, Aerts H, Yu S, et al. Development and validation of a prognostic model using blood biomarker information for prediction of survival of non-small-cell lung cancer patients treated with combined chemotherapy and radiation or radiotherapy alone (NCT00181519, NCT00573040, and NCT00572325). *Int J Radiat Oncol Biol Phys* 2011;81:360-368.

38. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441-446.

39. Lambin P, Petit SF, Aerts HJ, et al. The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: Exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol* 2010;96:145-152.

40. Deasy JO, Bentzen SM, Jackson A, et al. Improving normal tissue complication probability models: The need to adopt a "data-pooling" culture. *Int J Radiat Oncol Biol Phys* 2010;76:S151-S154.

41. Roelofs E, Persoon L, Nijsten S, et al. Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. *Radiother Oncol* 2013;108:174-179.