Radiotherapy and Oncology 136 (2019) 176–184

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Development and internal validation of a multinomial NTCP model for the severity of acute dyspnea after radiotherapy for lung cancer

Gilles Defraene [a,*], Ewoud Schuit [b], Dirk De Ruysscher [a,c]

[a] KU Leuven – University of Leuven, Department of Oncology, Experimental Radiation Oncology, Belgium; [b] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University; and [c] Maastricht University Medical Center, Department of Radiation Oncology (MAASTRO Clinic), GROW-School for Oncology and Developmental Biology, the Netherlands

ARTICLE INFO

Article history:
Received 13 October 2018
Received in revised form 22 February 2019
Accepted 29 March 2019
Available online 20 April 2019

Keywords:
Radiation-induced dyspnea
Patient-reported outcome
NTCP modeling
Lung cancer
Baseline symptoms

ABSTRACT

Background and purpose: Dyspnea evolution after radiotherapy for lung cancer is complex with potential symptom deterioration and improvement from baseline. We developed and internally validated a multinomial normal tissue complication probability (NTCP) model predicting dyspnea grade.

Materials and methods: Patient-reported dyspnea was collected pre-treatment and during 6 months follow-up for 182 stage I–IV lung cancer patients treated with radical (chemo)radiotherapy. Dyspnea changes (ΔDys) from the baseline grade (Dys_0) to the follow-up grade (Dys) were evaluated. A multinomial logistic regression model simultaneously predicting 3 grades of Dys (Dys $\geq$ 3, Dys = 2 and Dys $\leq$ 1 (reference level)) was optimized. Reference NTCP models predicting Dys $\geq$ 2 and Dys $\geq$ 3 risks irrespective of Dys_0 were generated for comparison. Models were shrunken and performance was assessed using optimism-corrected AUC (bootstrapping).

Results: Rates of ΔDys $\geq$ 1 (deterioration) and ΔDys $\leq$ -1 (improvement) at 6 months were 31.9% and 12.6%. Dys $\geq$ 3, Dys = 2 and Dys $\leq$ 1 rates were 13.7%, 20.9% and 65.4%, respectively. The multinomial model (combining the risk factors Dys_0 and MLD and the protective factor chemotherapy treatment) predicted Dys $\geq$ 3, Dys = 2 and Dys $\leq$ 1 with AUC (95% CI) of 0.72 (0.65–0.75) 0.76 (0.72–0.79) and 0.78 (0.74–0.80), respectively. Reference Dys $\geq$ 2 and Dys $\geq$ 3 models showed worse AUC: 0.64 (0.59–0.67) and 0.66 (0.50–0.70), respectively.

Conclusions: Dyspnea grade could be predicted with high accuracy using a multinomial NTCP model, yielding personalized dyspnea symptom improvement and deterioration risks.

 Radiotherapy and Oncology 136 (2019) 176–184

Radiation-induced pulmonary toxicity remains a major concern in the standard of care radical (chemo)radiotherapy treatment for lung cancer. Approximately 15% of patients suffer from severe symptoms in the first 6 months after treatment, while a larger subgroup present mild symptoms [1]. To manage the toxicity risk in clinical routine, normal tissue complication probability (NTCP) models predicting severe adverse events are applied. These models are based on the correlation of radiation pneumonitis with the mean lung dose (MLD) or the relative lung volume receiving 20 Gy (V_{20}) [2,3]. A MLD planning constraint of 20 Gy was recommended [4–6]. Upgraded MLD models have been proposed using relevant clinical factors associated with toxicity risk (e.g. age and smoking status) [6–9]. However, at best, moderate model accuracy (typically AUC < 0.65) with limited positive predictive value was reported [2–5]. Applying the models thus limits the severe pulmonary toxicity rate to 15%, but does not allow to reliably discriminate high risk from low risk patients before treatment. The consequence is that the 85% unaffected patients might have received a higher tumor dose leading to a possible gain in local control [10].

A more accurate prediction of pulmonary toxicity, possibly including mild symptoms, could thus allow an advanced treatment individualization (e.g. using adapted dose constraints for a subset of patients at highest risk of toxicity) and be beneficial to more patients. Other options for individualizing the treatment for selected high-risk patients may include altered radiotherapy treatments using conformal avoidance planning [11,12] and new techniques as proton therapy [13].

Multiple reasons exist for the poor accuracy and difficult external validation of NTCP models of pulmonary toxicity. Firstly, radiation-induced pulmonary toxicity is a highly multifactorial process involving many factors still to be elucidated [14]. Secondly, toxicity assessments are prone to observer dependence, even when they are scored according to published scales [15]. Regarding the latter, the feasibility of patient-reported outcome (PRO) tools was

* Corresponding author at: KU Leuven – University of Leuven, Department of Oncology, Experimental Radiation Oncology, B-3000 Leuven, Belgium.
E-mail address: gilles.defraene@uzleuven.be (G. Defraene).

https://doi.org/10.1016/j.radonc.2019.03.034
0167-8140/

shown, with the advantage of a better correlation to quality-of-life and a higher severity reported by patients than clinicians [16].

Another improvement could originate from a baseline symptom grading, which is usually not incorporated in prognostic modeling of pulmonary toxicity. Generally, approximately 10% of patients have severe symptoms at baseline, before treatment [1]. Reported risk factors for severe post-treatment symptoms could thus simply relate to the baseline symptoms of the patient, instead of reflecting a radiation-induced risk. Moreover, in a study on clinically relevant dyspnea (shortness of breath), almost 20% of patients with baseline symptoms were scored with less dyspnea after treatment [17]. A baseline symptom grade variable is thus likely to improve prognostic model accuracy. This way, the assessment of the incremental benefit from the introduction of new techniques could also be optimized, which is a prerequisite to optimize their cost-effective deployment [18].

In this work, prognostic factors of a PRO dyspnea grade change from baseline were identified. Consecutively, a multinomial logistic regression NTCP model simultaneously predicting different dyspnea outcome grades was developed. Prognostic factors of lung toxicity derived by Appelt et al. from a literature meta-analysis (MLD and 6 clinical factors [8,9]) were included and complemented by other published risk factors. The advantage over the classical NTCP modeling approach predicting the risk of severe follow-up dyspnea above a certain threshold grade (and without incorporating a baseline symptom grade variable) was assessed.

Materials and methods

Dataset

TRIPOD reporting guidelines were followed throughout this work [19]. Stage I–IV lung cancer patients treated with radical (chemo)radiotherapy between January 2014 and January 2017 at MAASTRO clinic were included. Exclusion criteria were stereotactic fractionated radiotherapy and a previous radiotherapy treatment in the thorax region. Treatment plans were calculated on a 50% expiration 4DCT frame or a mid-ventilation CT-scan [20] with Acuros (Eclipse, Varian Medical Systems, Palo Alto, CA) using volumetric modulated arc therapy (VMAT) or hybrid VMAT (VMAT component combined with an anteroposterior open-field component [21]) techniques. Margins of 5 mm from GTV to CTV and 5–10 mm from CTV to PTV were applied. Prescription doses were 66 Gy (2.75 Gy fractions with or without sequential chemotherapy or 2 Gy fractions with concurrent chemotherapy), 72 Gy (1.8 Gy fractions twice per day), 45 Gy (1.5 Gy fractions twice per day) possibly followed by 24 Gy (2 Gy fractions), or a dose escalation treatment up to 106.4 Gy (24 fractions, PET-boost randomized study [22]), with or without chemotherapy.

Patient-reported outcome (PRO) questionnaires were prospectively collected after written informed consent. The dyspnea question used in this work, was a vulgarized version of the Common Terminology Criteria for Adverse Events (CTCAE 4.0) question. It is reproduced in Appendix 1 in Dutch, together with a validated English version used in the REQUITE study [23]. PROs were collected during the last two weeks before the start of radiotherapy and subsequently at five different timepoints: 2 weeks after start of treatment, at the end of treatment, and at 3 weeks, 3 months and 6 months after the end of treatment. The 182 patients analyzed further returned questionnaires at the baseline timepoint and at least one follow-up timepoint beyond the end of treatment.

The clinical variables described by Appelt et al. (for grade ≥ 3 radiation pneumonitis) were retrospectively retrieved: age at start of treatment, chemotherapy regime (none, sequential or concurrent), tumor location (upper versus middle/lower lobe), current smoker (yes/no), history of smoking (yes/no) and pre-existing pulmonary comorbidity (yes/no) defined as chronic obstructive pulmonary disease (COPD) or other pulmonary disease [9]. As opposed to Appelt et al., age was analyzed as a continuous variable [7] and chemotherapy regime as concurrent/sequential versus none (few sequential chemotherapy cases in our dataset). Additionally, baseline World Health Organization performance status (WHO PS) [7], the prescription dose in EQD2,T (equivalent dose in 2 Gy fractions with overall treatment time (OTT) correction using $\alpha/\beta = 10$ Gy, a proliferation rate of 0.42 Gy/day and a reference time of 44 days, which is the median OTT of the 33 fraction treatments in our dataset [24]) and the planning target volume (PTV) [25] were included in the analyses. Dosimetric (physical dose) variables were the MLD (based on both lungs) [9], the relative total lung volume receiving more than 5 Gy (V_5) [26], the mean heart dose (MHD) and the maximal heart dose [27].

Logistic regression modeling and statistics

For every patient, the change in dyspnea grade from baseline (ΔDys) was defined as the difference between the maximal dyspnea grade scored during the 6 months follow-up (Dys) and the baseline dyspnea grade before treatment (Dys_0), with $\Delta Dys = Dys - Dys_0$. The association of Dys and ΔDys with dosimetric features was assessed using linear regression.

Univariable logistic regression was performed for 4 endpoints: any worsening in dyspnea grade with respect to the baseline grade ($\Delta Dys \geq 1$), any improvement in dyspnea grade with respect to the baseline grade ($\Delta Dys \leq -1$), crude dyspnea of grade 2 or more ($Dys \geq 2$) and grade 3 or more ($Dys \geq 3$). Patients with $Dys_0 = 0$ were excluded for the $\Delta Dys \leq -1$ endpoint analysis as their symptoms could by definition not improve. For the continuous variables, non-linear associations with outcome were tested using restricted cubic splines. No missing data were present except for the heart dose variables of 8 patients for which single imputation was performed based on their PTV size and lung dose.

For the $Dys \geq 2$ and $Dys \geq 3$ endpoints, variables with univariable $p < 0.3$ were included in a multivariable logistic regression modeling. For both endpoints, two separate modeling efforts were conducted: one excluding and one including the Dys_0 variable in order to assess its impact on classical NTCP modeling performance. Further, a multinomial logistic regression model was optimized using all variables with $p < 0.3$ in univariable multinomial logistic regression. For all modeling efforts, the best (lowest p value) lung and heart dose variable were added to the variable list (if none were included otherwise). Best subsets (multinomial) logistic regression was performed based on the Akaike information criterion (AIC), which incorporates an assessment of the likelihood function and a penalty on the number of model parameters to avoid overfitting. Model calibration was assessed by calibration plots and the Hosmer–Lemeshow test, where $p < 0.05$ indicates a bad model fit to the data points [28]. Model discrimination was assessed using the discrimination slope (average prediction difference between patients with and without the outcome) and the AUC of the receiver operating characteristic (ROC) curve. For the multinomial model 3 AUCs were calculated, one for each outcome level ($Dys \geq 3$, $Dys = 2$ and $Dys \leq 1$) [29]. The reported models were optimized on the complete dataset. Internal validation was performed using 100 bootstrap samples (obtained from the original dataset allowing replacement by sampling the same individual multiple times). Within each bootstrap sample, all NTCP modeling steps were repeated. This resulted in a shrinkage factor to be multiplied with the regression coefficients. Using the thus obtained shrunk coefficients should give more generalizable predictions out of sample [28]. Reported AUCs were corrected for optimism by averaging the 100 AUCs obtained when applying bootstrap sample-derived model coefficients to predict risks in the whole dataset [28].

Differences in model predictions between the Dys $\geq$ 3 models, the multinomial model and the reference Appelt model [9] were assessed per Dys_0 grade, and for 3 risk categories of patients. A low-risk, average-risk and high-risk patient was simulated for each Dys_0 category. The hypothetical patients' characteristics were derived from our dataset as respectively the 25th, 50th and 75th percentile value for every risk factor, within the subgroup of patients with the same Dys_0 grade. All statistics and model building were performed in Statistica version 13 (Dell Inc., Tulsa, OK) and Matlab R2015b (The Mathworks Inc., Natick, MA).

Results

Patient and treatment characteristics are listed in Table 1. The distribution of dyspnea grades at baseline and follow-up are depicted in Fig. 1a. 55.5% of patients reported a stable maximal dyspnea grade at 6 months, while 31.9% worsened and 12.6% improved. A significant gradual increase in MLD ($p = 0.033$) and lung V_5 ($p = 0.044$) was observed between subgroups of patients with ΔDys = -1 (1 grade improvement, MLD = 11.5 Gy and V_5 = 53.9% on average) up to ΔDys = 3 (3 grades worsening, MLD = 16.5 Gy and V_5 = 76.1% on average), Fig. 1b. Similar dependency could not be observed for heart dose metrics, nor for subgroups based on crude follow-up dyspnea grades.

In univariable logistic regression, better WHO PS, lower Dys_0 and no pulmonary comorbidity were significantly associated with $\Delta Dys \geq 1$ (Appendix 2). Higher Dys_0, upper lobe tumor location and lower prescription dose were predictors for $\Delta D \leq -1$. MLD and lung V_5 were borderline significant for $\Delta Dys \geq 1$ ($p < 0.1$). In univariable multinomial logistic regression, significant associations were found for Dys_0, WHO PS and chemotherapy regime (Table 2). The same variables were significant in univariable logistic regression for Dys $\geq$ 3 and/or Dys $\geq$ 2. There was no evidence for transforming any continuous variable.

The multivariable model coefficients and performance metrics are reported in Table 3. The optimism-corrected AUC was 0.78 (0.70–0.80) and 0.70 (0.55–0.74) for Dys $\geq$ 2 and Dys $\geq$ 3, respectively. Excluding Dys_0 from the models resulted in significantly worse fits (Likelihood ratio test $p < 0.01$, AUC = 0.64–0.66). Note that a validation of the Appelt model in our dataset resulted in even worse discrimination, AUC = 0.62 (0.50–0.74), and calibration (Appendix 5). The optimal multinomial model included the covariates Dys_0, MLD and chemotherapy regime and had AUC of 0.72 (0.65–0.75), 0.76 (0.72–0.79) and 0.78 (0.74–0.80) for the prediction of Dys $\geq$ 3, Dys = 2 and Dys $\leq$ 1, respectively. Discrimination slopes of these predictions were 0.11, 0.15 and 0.27, respectively. Fig. 2 depicts NTCP model behavior for selected subgroups of patients according to the model variables. All models combined MLD with some clinical factors. From the multinomial model, the probability of worsening and improvement of dyspnea grade can be calculated, e.g. for a patient with Dys_0 = 2 at MLD = 20 Gy, the probabilities of Dys $\leq$ 1 (improvement), Dys = 2 (stable) and Dys $\geq$ 3 (worsening) are with chemotherapy (without chemotherapy between brackets): 25.5% (12.4%), 32.7% (35.6%) and 41.9 % (52.0%), respectively. Model coefficients, performance metrics and graphs to be used with an MLD corrected for fractionation and OTT (EQD2,T) instead of an MLD in physical dose are reported in Appendix 6 for all the presented models.

Our classical Dys $\geq$ 3 model excluding Dys_0 resulted in similar probabilities as the Appelt model in low- and average-risk patients (Fig. 3). For high-risk patients with Dys_0 = 0, an absolute 12% increased risk prediction was observed with the Appelt model which includes 6 clinical risk factors (compared to only 2 clinical factors in the Dys $\geq$ 3 model). The differences with the model predictions including Dys_0 were larger.

Table 1
Patient and treatment characteristics in the dataset of 182 patients. For dosimetry, physical doses are reported in Gy. Median and range or absolute numbers and proportions.

Age (years)	67 (44–88)
Tumor T stage	
0,1,X	28 (15.4%)
2	52 (28.6%)
3	29 (15.9%)
4	43 (23.6%)
Unknown	30 (16.5%)
N stage	
0,1,X	56 (30.8%)
2	58 (31.9%)
3	38 (20.9%)
Unknown	30 (16.5%)
Histology	
Squamous cell carcinoma	45 (24.7%)
Adenocarcinoma	44 (24.2%)
NSCLC NOS	34 (18.7%)
Small cell	28 (15.4%)
Other	12 (6.6%)
Unknown	19 (10.4%)
WHO performance status at baseline	
0	40 (22.0%)
1	122 (67.0%)
2	15 (8.2%)
3	5 (2.7%)
Pre-existing pulmonary comorbidity	
Yes	97 (53.3%)
No	75 (41.2%)
Smoking status	
Current smoker	43 (23.6%)
Stopped	133 (73.1%)
Never smoker	6 (3.3%)
Chemotherapy treatment	
Concurrent	99 (54.4%)
Sequential	13 (7.1%)
None	70 (38.5%)
Lobe location	
Upper	106 (58.2 %)
Middle	15 (8.2%)
Lower	56 (30.8%)
Combined	2 (1.1%)
Tx	3 (1.6%)
Treatment technique	
VMAT	148 (81.3%)
Hybrid VMAT	34 (18.7%)
PTV volume (cc)	286.9 (24.1–1948.8)
Dose per fraction (Gy)	
1.5	81 (44.5%)
1.8	22 (12.1%)
2	22 (12.1%)
2.75	50 (27.5%)
$\geq$ 3	7 (3.9%)
Mean Lung Dose (Gy)	13.2 (1.9–23.4)
Lung V_5 (%)	56.5 (8.6–100.0)
Mean Heart Dose (Gy)	6.1 (0.1–34.1)
Heart D_{max} (Gy)	60.5 (0.3–78.6)

Abbreviations: NSCLC NOS: Non-small cell lung cancer Not otherwise specified; WHO: World Health Organization; VMAT: Volumetric Modulated Arc Therapy; PTV: Planning Target Volume; Lung V_5: Relative total lung volume receiving at least 5 Gy.

Discussion

Accurate prognostic models of pulmonary toxicity are crucial to personalize the definitive chemoradiotherapy treatment for lung cancer. In this work, we optimized NTCP models for patient-reported acute dyspnea in the first 6 months after VMAT treatment, starting from a set of published clinical and dosimetric prognostic factors. The inclusion of the baseline dyspnea grade (Dys_0) had some important advantages. First, a multinomial NTCP model including the Dys_0 covariate could simultaneously predict the

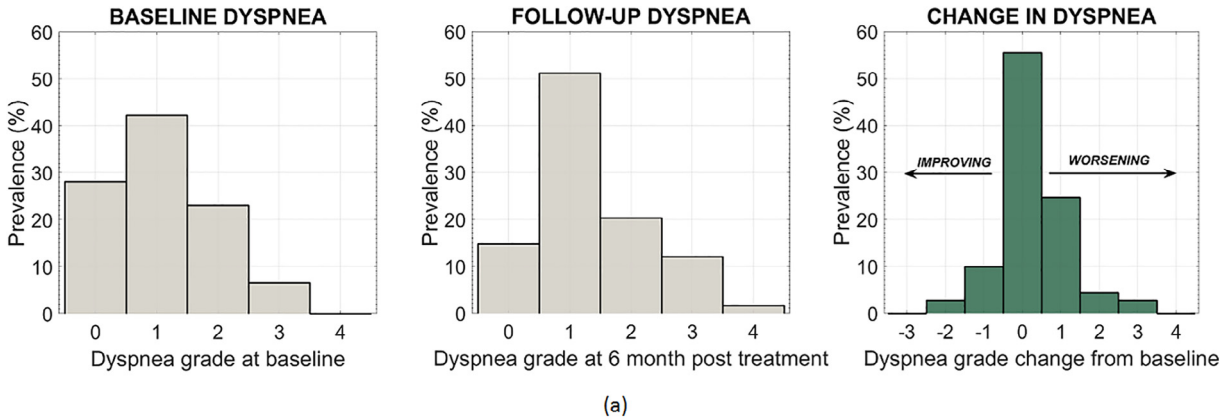

Fig. 1. (a) Baseline (Dys^0) *(left)* and 6 months follow-up (Dys) *(central)* dyspnea grade distributions in the dataset of 182 patient-reported outcomes. Change in dyspnea grade from baseline (ΔDys) *(right)* with positive values indicating a number of grades worsening of dyspnea symptoms and negative values indicating improving dyspnea symptoms by a number of grades. (b) Mean Lung Dose (MLD) *(upper row)* and relative lung volume receiving 5 Gy (V5) *(lower row)* for different categories of crude follow-up dyspnea grade (Dys) *(left column)* and dyspnea grade change from baseline (ΔDys) *(right column)*. Mean values, with 1 and 2 standard errors for box and whisker positions. The number of patients per category is shown on top of the bars. For ΔDys, a significant lung dose increase is observed between the grades −1 and + 3 (containing 97.8% of patients), both for MLD and V5 ($p = 0.033$ and $p = 0.044$, respectively). For the crude follow-up dyspnea, the lung dose dependence is more flat, with increased MLD only observed between grade 2 and 4 (involving only 34.6% of patients and $p > 0.05$). Physical lung doses are reported.

grade ≤ 1, grade = 2 and grade ≥ 3 dyspnea probabilities at 6 months with high accuracy. We also observed a considerably higher discriminative power when including the knowledge on Dys_0 in classical NTCP models predicting Dys ≥ 2 and Dys ≥ 3 endpoints ($p < 0.01$). Secondly, prognostic factors for a 'change in dyspnea grade from baseline' (ΔDys) could be analyzed. Worsening dyspnea symptoms by at least 1 grade were associated to a better performance status, a lower Dys_0 and no pulmonary comorbidity. A major advantage of this approach was the potential to also study associations with improvement in dyspnea grade, which is regularly observed after treatment (13% of patients in our dataset). We showed that upper lobe tumors, a higher Dys_0 and lower pre-

Table 2
Univariable (multinomial) logistic regression results (OR, *p* value and AUC) for 13 variables. The upper half of the table shows two classical follow-up dyspnea endpoints predicting severe dyspnea above a threshold grade. The lower half of the table shows a multinomial dyspnea grade endpoint simultaneously predicting 3 levels of dyspnea (the two columns present model coefficients and *p* values for the prediction of two outcome levels with respect to the reference level). Significant associations ($p < 0.05$) are shown in bold.

Classical follow-up dyspnea (Dys) prediction	Dys $\geq$ 2 endpoint			Dys $\geq$ 3 endpoint		
	OR (95% CI)	*p*	AUC	OR (95% CI)	*p*	AUC
MLD (+1 Gy)	1.009 (0.950; 1.071)	0.78	0.51	1.076 (0.986; 1.175)	0.10	0.60
Lung V_5 (+1%)	0.905 (0.266; 3.077)	0.87	0.51	2.595 (0.471; 14.313)	0.27	0.57
MHD (+1 Gy)	0.995 (0.954; 1.036)	0.80	0.48	1.018 (0.964; 1.075)	0.53	0.55
Heart D_{max} (+1 Gy)	1.003 (0.991; 1.016)	0.63	0.52	1.012 (0.992; 1.032)	0.24	0.59
Prescription dose (Gy EQD2, T)	1.013 (0.986; 1.040)	0.35	0.59	1.006 (0.971; 1.043)	0.72	0.57
Age (+1 year)	1.015 (0.982; 1.049)	0.39	0.54	1.044 (0.995; 1.095)	0.078	0.61
PTV (+1 cc)	1.000 (0.999; 1.001)	0.93	0.52	1.001 (1.000; 1.002)	0.22	0.60
WHO PS (reference = 0)		**0.0036**	0.62		0.075	0.61
1	2.659 (1.086; 6.511)	0.999		3.288 (0.728; 14.847)	0.052	
$\geq$2	7.071 (2.107; 23.728)	**0.0041**		6.333 (1.106; 36.276)	0.60	
Upper lobe (vs middle/lower)	1.086 (0.579; 2.037)	0.80	0.51	1.495 (0.608; 3.673)	0.38	0.55
Pulmonary comorbidity (vs no)	1.783 (0.951; 3.343)	0.071	0.57	1.646 (0.686; 3.949)	0.26	0.56
Chemotherapy (vs none)	0.484 (0.259; 0.906)	**0.023**	0.59	0.746 (0.318; 1.751)	0.50	0.54
Current smoker (vs no)	0.887 (0.429; 1.834)	0.75	0.51	0.783 (0.275; 2.228)	0.65	0.52
Baseline dyspnea (reference = 0)		**<0.001**	0.76		**0.0037**	0.67
$Dys_0 = 1$	2.402 (0.886; 6.517)	0.085		0.972 (0.260; 3.632)	0.97	
$Dys_0 \geq 2$	16.392 (5.870; 45.775)	**<0.001**		4.313 (1.323; 14.056)	**0.015**	
Multinomial follow-up dyspnea grade (Dys $\leq$ 1, Dys = 2 and Dys $\geq$ 3) predictions	Dys = 2 (versus Dys $\leq$ 1)			Dys $\geq$ 3 (versus Dys $\leq$ 1)		
	Coefficient (95% CI)		*p*	Coefficient (95% CI)		*p*
MLD (+1 Gy)	−0.027 (−0.099; 0.044)		0.45	0.067 (−0.023; 0.157)		0.14
Lung V_5 (+1%)	−0.69 (−2.17; 0.79)		0.36	0.79 (−0.95; 2.53)		0.37
MHD (+1 Gy)	−0.019 (−0.070; 0.033)		0.48	0.013 (−0.042; 0.069)		0.64
Heart D_{max} (+1 Gy)	−0.0013 (−0.016; 0.013)		0.86	0.011 (−0.0084; 0.031)		0.26
Prescription dose (Gy EQD2,T)	0.014 (−0.017; 0.045)		0.37	0.010 (−0.027; 0.047)		0.59
Age (+1 year)	−0.0026 (0.042; 0.037)		0.90	0.042 (−0.0064; 0.091)		0.089
PTV (+1 cc)	−0.00068 (−0.0021; 0.00073)		0.34	0.00063 (−0.00062; 0.0019)		0.32
WHO PS (reference = 0)						
1	0.79 (−0.25; 1.83)		0.14	1.34 (−0.18; 2.85)		0.084
$\geq$2	1.75 (0.37; 3.14)		**0.013**	2.33 (0.52; 4.15)		**0.012**
Upper lobe (vs middle/lower)	−0.11 (−0.85; 0.64)		0.78	0.38 (−0.54; 1.29)		0.42
Pulmonary comorbidity (vs no)	0.55 (−0.21; 1.30)		0.16	0.63 (−0.27; 1.52)		0.17
Chemotherapy (vs none)	−0.86 (−1.61; −0.12)		**0.023**	−0.52 (−1.39; 0.36)		0.25
Current smoker (vs no)	−0.038 (−0.89; 0.82)		0.93	−0.25 (−1.32; 0.81)		0.64
Baseline dyspnea (reference = 0)						
$Dys_0 = 1$	1.60 (0.056; 3.13)		**0.042**	0.13 (−1.20; 1.45)		0.85
$Dys_0 \geq 2$	3.39 (1.84; 4.94)		**<0.001**	2.27 (1.036; 3.51)		**<0.001**

Abbreviations: MLD: Mean Lung Dose; MHD: Mean Heart Dose; Lung V_5: Relative total lung volume receiving at least 5 Gy; EQD2,T: Equivalent dose in 2 Gy fractions with time correction; PTV: Planning Target Volume; WHO PS: World Health Organization Performance Status; OR: Odds Ratio; CI: Confidence Interval; AUC: Area Under the Curve; vs: versus.

scription doses were univariably associated with a higher probability of improving symptoms at 6 months. Dyspnea symptom changes also allowed to detect a significant lung dose response, as gradually increasing lung doses were associated with subgroups of patients, from the 1 grade improvement subgroup up to the 3 grades worsening subgroup. Only the five outlier patients reporting two grades dyspnea improvement did not receive the lowest lung doses (Fig. 1b). These were all upper lobe patients with COPD and a very large PTV (median of 799 cc, compared to 282 cc in the rest of the population). Shrinkage of tumor volume and COPD lung bullae are common phenomena after radiotherapy and could be an explanation of the strong dyspnea improvement in these patients.

The multinomial NTCP model indicates a strong MLD dependence of the Dys $\geq$ 3 risk at 6 months, similarly to the classical Dys $\geq$ 3 NTCP model. Limiting MLD is clearly beneficial in that light, but additionally, it also increases the probability of Dys $\leq$ 1 in selected subgroups of patients.

Clinical factors were found to be at least as important as dosimetry in several lung toxicity studies [7–9]. Interestingly, our initial Dys $\geq$ 2 and Dys $\geq$ 3 models included with WHO PS and age two identical risk factors as the model of Dehing et al. predicting physician-rated acute dyspnea grade $\geq$ 2 (AUC = 0.62) [7]. The 2 other clinical variables in their model, current smoking and forced expiratory volume, could not be validated (OR < 1 but not significant) or were not assessed, respectively. Dys_0 replaced WHO PS in our final Dys $\geq$ 2 and Dys $\geq$ 3 models and significantly improved AUC. Interestingly, Dys_0 was also a protective factor for $\Delta Dys \geq 1$ indicating a lower risk of worsening when starting from a higher baseline dyspnea grade, as well as a prognostic factor for $\Delta Dys \leq -1$ (symptom improvement). For the other risk factors reported by Appelt et al., an association of middle/lower lobe location was found with non-improving symptoms in our work, rather than being a risk factor for severe symptoms as was repeatedly reported [9,30–32]. Pulmonary comorbidity was univariably associated to a dyspnea worsening in our work, but as a protective factor [9,33]. Chemotherapy was only significantly associated univariably to a crude endpoint ($p = 0.02$ for Dys $\geq$ 2), and it was selected in the multinomial model, both as a protective factor. As no significant association of chemotherapy (versus no chemotherapy) was found for acute radiation pneumonitis in reported meta-analyses, the protective effect in our dataset is most likely due to selection of the fittest patients for chemotherapy rather than a sensitizing effect of chemotherapy [34]. As dosimetric risk factor, no final choice could be made based on our data between MLD and lung V_5. MLD was selected in the multinomial model, however, a model including lung V_5 performed similarly (third best AIC

Table 3

(upper part) Multivariable logistic regression results for the 2 classical follow-up dyspnea endpoints. Models were optimized without (the approach in literature) and with the baseline dyspnea grade (Dys_0) variable. Covariates (coefficients and OR), discriminative power and calibration for the best subsets logistic regression model selected based on AIC. Note the large improvement of discrimination (AUC and slope) when including the knowledge on Dys_0. *(lower part)* Multinomial logistic regression model predicting 3 categories of follow-up dyspnea ($Dys \leq 1$, $Dys = 2$ and $Dys \geq 3$) simultaneously. Calibration plots of all models are depicted in Appendix 3. Internal validation using bootstrapping resulted in shrinkage factors (f) for the model coefficients and optimism-corrected AUC values.

Classical follow-up dyspnea (Dys) prediction	Model coefficient		Apparent OR (95% CI)	*p*	AUC (95% CI)	Discr. slope	Calibr. HL *p* value
Without 'baseline dyspnea' (Dys_0) variable							
$Dys \geq 2$ endpoint	Apparent	Shrunk (f = 0.76)			**Apparent:**	**0.087**	**0.38**
Intercept	−1.788	−1.359			**0.671 (0.586; 0.758)**		
WHO PS = 1 (vs 0)	0.944	0.717	2.571 (1.039; 6.360)	0.041	**Optimism-corrected:**		
WHO PS ≥ 2 (vs 0)	1.846	1.403	6.338 (1.819; 22.081)	0.0037	**0.641 (0.586; 0.673)**		
MLD (+1 Gy)	0.058	0.044	1.059 (0.985; 1.139)	0.12			
Chemotherapy (vs none)	−0.824	−0.626	0.439 (0.209; 0.920)	0.029			
$Dys \geq 3$ endpoint	Apparent	Shrunk (f = 0.66)			**Apparent:**	**0.041**	**0.27**
Intercept	−6.270	−3.725			**0.634 (0.510; 0.758)**		
MLD (+1 Gy)	0.083	0.055	1.086 (0.993; 1.189)	0.069	**Optimism-corrected:**		
Age (+1 year)	0.048	0.036	1.049 (0.999; 1.102)	0.054	**0.655 (0.500; 0.699)**		
With 'baseline dyspnea' (Dys_0) variable							
$Dys \geq 2$ endpoint	Apparent	Shrunk (f = 0.86)			**Apparent:**	**0.268**	**0.34**
Intercept	−2.650	−2.248			**0.792 (0.718; 0.865)**		
$Dys_0 = 1$ (vs 0)	0.977	0.840	2.656 (0.958; 7.363)	0.060	**Optimism-Corrected:**		
$Dys_0 \geq 2$ (vs 0)	2.883	2.479	17.876 (6.143; 52.019)	<0.001	**0.776 (0.698; 0.802)**		
MLD (+1 Gy)	0.082	0.071	1.086 (0.999; 1.180)	0.051			
Chemotherapy (vs none)	−0.859	−0.739	0.424 (0.184; 0.976)	0.044			
$Dys \geq 3$ endpoint	Apparent	Shrunk (f = 0.61)			**Apparent:**	**0.122**	**0.19**
Intercept	−7.036	−4.292			**0.744 (0.628; 0.860)**		
$Dys_0 = 1$ (vs 0)	0.121	0.074	1.129 (0.294; 4.331)	0.86	**Optimism-corrected:**		
$Dys_0 \geq 2$ (vs 0)	1.659	1.012	5.255 (1.526; 18.101)	0.0086	**0.696 (0.550; 0.742)**		
MLD (+1 Gy)	0.109	0.066	1.115 (1.011; 1.230)	0.029			
Age (+1 year)	0.043	0.026	1.044 (0.990; 1.101)	0.11			
Multinomial follow-up dyspnea grade ($Dys \leq 1$, $Dys = 2$, $Dys \geq 3$) predictions	Model coefficient		95% CI of apparent model coefficient	*p*	AUC (95% CI)	Discr. slope	Calibr. HL *p* value
$Dys = 2$ (versus $Dys \leq 1$)	Apparent	Shrunk (f = 0.78)			***Dys ≤ 1***	**0.266**	**0.36**
Intercept	−3.186	−2.485	(−5.044; −1.327)	<0.001	**Apparent:**		
MLD (+1 Gy)	0.043	0.034	(−0.051; 0.137)	0.37	**0.790 (0.726; 0.855)**		
Chemotherapy (vs none)	−0.807	−0.629	(−1.755; 0.140)	0.095	**Optimism-corrected:**		
$Dys_0 = 1$ (vs 0)	1.635	1.275	(0.0830; 3.186)	0.039	**0.778 (0.735; 0.802)**		
$Dys_0 \geq 2$ (vs 0)	3.382	2.638	(1.810; 4.954)	<0.001	***Dys = 2***	**0.150**	**0.40**
$Dys \geq 3$ (versus $Dys \leq 1$)	Apparent	Shrunk (f = 0.76)			**Apparent:**		
Intercept	−4.081	−3.102	(−6.082; −2.081)	<0.001	**0.761 (0.667; 0.856)**		
MLD (+1 Gy)	0.150	0.114	(0.031; 0.269)	0.014	**Optimism-corrected:**		
Chemotherapy (vs none)	−0.937	−0.712	(−2.059; 0.184)	0.10	**0.763 (0.723; 0.787)**		
$Dys_0 = 1$ (vs 0)	0.299	0.227	(−1.056; 1.655)	0.67	***Dys ≥ 3***	**0.106**	**0.81**
$Dys_0 \geq 2$ (vs 0)	2.516	1.912	(1.209; 3.822)	<0.001	**Apparent:**		
					0.736 (0.619; 0.853)		
					Optimism-corrected:		
					0.716 (0.651; 0.748)		

Abbreviations: MLD: Mean Lung Dose; WHO PS: World Health Organization Performance Status; OR: Odds Ratio; Discr. slope: Discrimination slope; CI: Confidence Interval; AUC: Area Under the Curve; HL: Hosmer–Lemeshow test; Dys_0: baseline dyspnea grade; Dys: follow-up dyspnea grade.

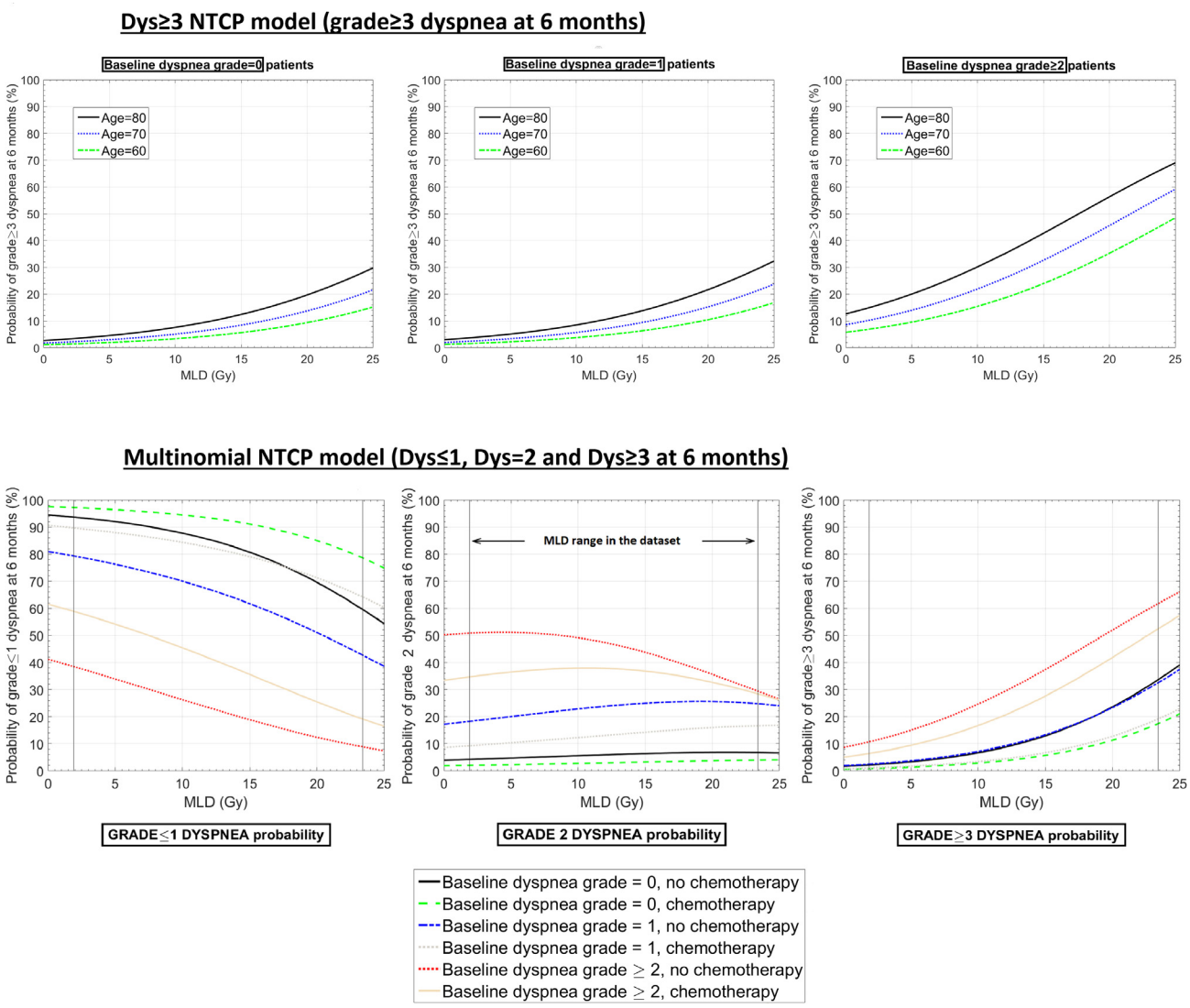

Fig. 2. Graphical representation of the Dys ≥ 3 and multinomial logistic regression NTCP models (apparent model coefficients, no shrinkage applied) respectively predicting grade ≥ 3 dyspnea, and simultaneously grade ≤ 1, grade = 2 and grade ≥ 3 dyspnea at 6 months after treatment. *Dys ≥ 3 model (upper row)*: Mean Lung Dose (MLD) dependence of the risks for representative patient categories based on the model variable age, for patients with $Dys^0 = 0$ *(left graph)*, $Dys^0 = 1$ *(central graph)* and $Dys^0 = \geq 2$ *(right graph)*. *Multinomial model (lower row)*: MLD dependence for representative patient categories based on the model variables Dys^0 and chemotherapy treatment for the risk of Dys ≤ 1 *(left graph)*, Dys = 2 *(central graph)* and Dys ≥ 3 *(right graph)* at 6 months. Predicted probabilities can be calculated using following formulae: for the Dys ≥ 3 model, $NTCP = (1 + e^{-S})^{-1}$ with $S = -7.036 + 0.121$ (if $Dys^0 = 1$) + 1.659 (if $Dys^0 \geq 2$) + 0.043*Age + 0.109*MLD; for the multinomial model, $NTCP(Dys \leq 1) = (1 + e^{S1} + e^{S2})^{-1}$, $NTCP(Dys = 2) = e^{S1}*(1 + e^{S1} + e^{S2})^{-1}$, $NTCP(Dys \geq 3) = e^{S2}*(1 + e^{S1} + e^{S2})^{-1}$, with $S1 = -3.186 + 0.043*MLD - 0.807$ (if Chemotherapy = yes) + 1.635 (if $Dys^0 = 1$) + 3.382 (if $Dys^0 \geq 2$) and $S2 = -4.081 + 0.150*MLD - 0.937$ (if Chemotherapy = yes) + 0.299 (if $Dys^0 = 1$) + 2.516 (if $Dys^0 \geq 2$). Appendix 4 depicts the Dys ≥ 2 model. MLD is reported in physical dose.

model). A very high range of V_5 was present in this VMAT dataset, including 16 patients with V_5 of 100% [26]. A low dose volume effect in pulmonary toxicity could thus not be excluded. Physical lung and heart doses were used in our analyses. In order to obtain models that can be used with biologically corrected MLD data, all presented models were refitted with MLD recalculated in equivalent dose in 2 Gy fractions, using alpha/beta = 4 Gy for fractionation correction and a dose recovery of 0.44 Gy/day for OTT correction as reported by Bentzen et al. [24]. Appendix 6 shows the coefficients, model behavior and performance metrics of these biologically corrected MLD-based NTCP models. A slightly lower optimism-corrected AUC and MLD dependence were generally observed for these models compared to the physical MLD-based NTCP models. For similar fractionation schedules as in this work, it is therefore suggested to use the physical MLD-based NTCP models. Studies on the association of heart dose and pulmonary toxicity resulted in conflicting reports [27,35]. No association to any dyspnea (change) endpoint was observed in our work. A dyspnea improvement after treatment was previously observed but not modeled [17]. Even though tumor shrinkage is likely involved in dyspnea improvement, it appeared that patients improving by 1 grade also formed the subgroup with the lowest MLD and lung V_5. This was independent from the finding of more dyspnea improvement in upper lobe locations (similar average MLD of 12.7 Gy and 13.1 Gy for upper and middle/lower lobe locations, respectively). Moreover, lower prescription doses were associated univariably to a dyspnea improvement (Appendix 2), contradicting the hypothesis of a dose-dependent tumor shrinkage impacting dyspnea in our dataset.

This study had some limitations that should be addressed. AIC was used to limit overfitting together with a non-ideal preselection of variables based on univariable associations, which resulted

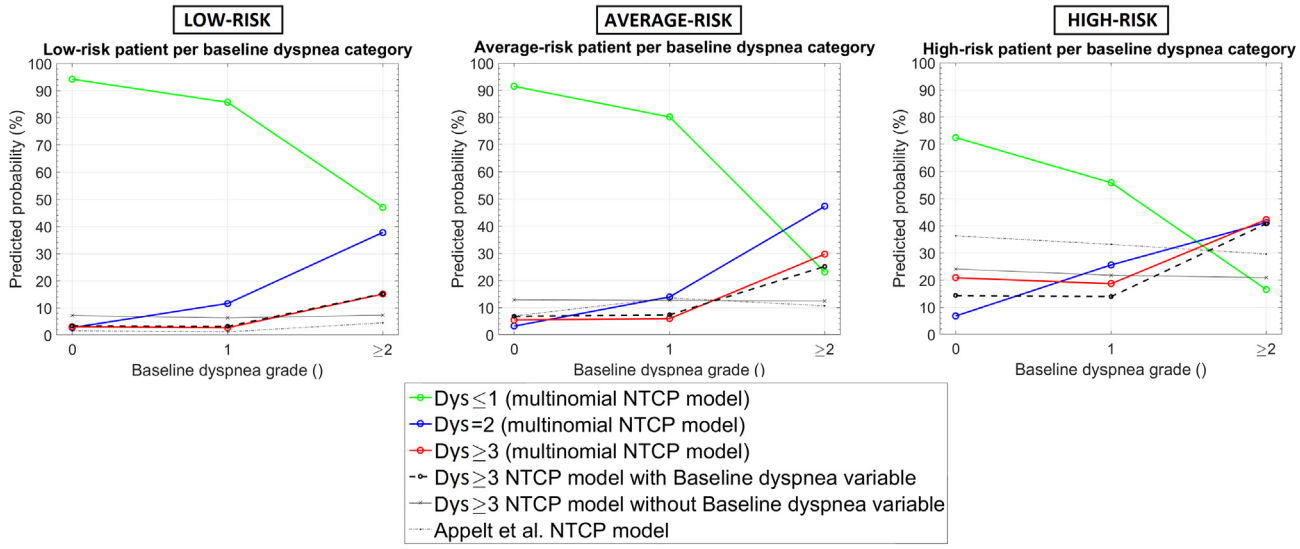

Fig. 3. Predictions according to the multinomial model, Dys ≥ 3 models (risk of grade ≥ 3 dyspnea) with and without the Dys^0 covariate, and the reference model published by Appelt et al. (risk of symptomatic (CTCAE 3.0 grade ≥ 3) radiation pneumonitis [9]) in function of the baseline dyspnea grade before treatment. Predicted probabilities for a hypothetical low-risk *(left graph)*, average-risk *(central graph)* and high-risk *(right graph)* patient representing each Dys^0 category. The hypothetical patients' characteristics were derived from our dataset as respectively the 25th, 50th and 75th percentile value for every risk factor, per subgroup of patients with the same Dys^0. Risk factors were MLD, chemotherapy and Dys^0 in the multinomial model; MLD, age and Dys^0 in the optimal Dys ≥ 3 model; MLD and age in the classical D ≥ 3 model (excluding Dys^0); MLD, age, pulmonary comorbidity, current smoker, former smoker, sequential chemotherapy and lower/middle lobe location in the Appelt model. Note the dependence of the multinomial and optimal Dys ≥ 3 model predictions on Dys^0 resulting in a broader range of predicted risks. This effect was not picked up by the classical models (more flat curves in function of Dys^0).

in a multinomial model with 8 degrees of freedom. External validation remains of paramount importance to confirm model exportability [5], also in light of the considerable model coefficient shrinkage factors (0.76–0.78) obtained through internal validation. However, the significant stepwise lung dose increase observed per dyspnea grade change subgroup, the optimism-corrected AUC 95% CI lower bound reaching 0.65–0.74 for the different outcome level predictions from the multinomial model and its excellent calibration (Appendix 3), robustly showed the advantage of the new approach including baseline dyspnea. Larger datasets are required to analyze all baseline and follow-up dyspnea grades separately (e.g. too few $Dys_0 = 3$ and Dys = 4 cases were currently available), which could result in a further refinement of the model.

Secondly, physical lung and heart doses were used in the presented models. As accelerated treatments are predominant in our dataset, updated models (coefficients, performance metrics and graphs) are presented for biologically corrected MLD in Appendix 6. Future studies could validate both the physical MLD-based and the biologically corrected MLD-based models in order to select the best performing model for a specific situation, and could search for more optimal dose correction parameters for the (multinomial) dyspnea endpoint.

Finally, not every patient scored dyspnea at all timepoints (only 28% of patients returned forms at all 4 post-treatment time points). We therefore used the maximal grade reported during the first 6 months after treatment. For approximately 24% of patients in our dataset, no later score than 2 months after treatment was available. This could have created a bias, as potential toxicity events at later time were unreported in these patients and the reason for the absence of later PRO data was unknown. To evaluate this potential bias, we conducted an additional modeling analysis for every endpoint, exclusively using the 139 patients having PRO data collected up to a later timepoint than 2 months. Except for pulmonary comorbidity the same variables as shown in Table 2 were univariaby significant (data not shown). Based on the same covariates of Table 3, newly optimized models resulted in similar AUC for all endpoints (difference in AUC < 0.03).

Moreover, some patients could have been lost to follow-up (not returning any follow-up questionnaire beyond the end of treatment) or could have died in the first months after the end of treatment. While no data on study compliance was available for our cohort, another observational study (REQUITE project [23]) using an identical routine of sending PRO by post to the patient, has shown that in lung cancer approximately 10% of patients did not return any of their follow-up forms, while less than 4% mortality was observed in that patient group in the first 3 months after the end of treatment. The expected rate of pulmonary toxicity-related deaths that were not included in our dataset should thus be very low and is unlikely to have severely impacted the reported model coefficients. Sending PRO by post should have avoided as much as possible any bias of patient status/illness interfering with study compliance. However, validation of our models in large prospective studies recording the cause of death for every patient is required. Such studies should also take into account information on progression status as this could have influenced worsening of dyspnea symptoms in some patients.

In conclusion, to the best of our knowledge, this is the first study modeling acute PRO dyspnea leading to a well discriminating multinomial NTCP model of dyspnea grade. For clinical application of the multinomial NTCP model, a possible approach could be to maximize the absolute difference between the predicted probabilities of dyspnea improvement and worsening (with respect to Dys_0). Using ΔDys could also highlight clinical factors relevant to the radiotherapy treatment and independent from Dys_0. The subtle lung dose dependence of ΔDys could also signify a step forward in understanding the role of imaging-derived lung damage [36–37]. These endpoints exhibited a strong dose dependence but could not be consistently associated to crude toxicity endpoints [36]. Finally, biophysics and radiogenomics have been discussed recently as potentially aiding lung toxicity prognosis and should be part of future modeling efforts [14,38].

Conflict of interest statement

Nothing to disclose.

Acknowledgement

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement 601826 (REQUITE).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2019.03.034.

References

[1] van Baardwijk A, Wanders S, Boersma L, Borger J, Öllers M, Dingemans A-M, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I-III non-small cell lung cancer. J Clin Oncol 2010;28:1380–6.
[2] Kwa SL, Lebesque JV, Theuws JC, et al. Radiation pneumonitis as a function of mean lung dose : an analysis of pooled data of 540 patients. Int J Radiat Oncol Biol Phys 1998;42:1–9.
[3] Seppenwoolde Y, Lebesque JV, De Jeager K, et al. Comparing different NTCP models that predict the incidence of radiation pneumonitis. Int J Radiat Oncol Biol Phys 2003;55:724–35.
[4] Marks LB, Bentzen SM, Deasy JO, et al. Radiation dose-volume effects in the lung. Int J Radiat Oncol Biol Phys 2010;76:S70–6.
[5] Bentzen S, Constine L, Deasy J, Eisbruch A, Jackson A, Marks L, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an introduction to the scientific issues. Int J Radiat Oncol Biol Phys 2010;76:S3–9.
[6] De Ruysscher D, Faivre-Finn C, Moeller D, Nestle U, Hurkmans C, Le Péchoux C, et al. European Organization for Research and Training of Cancer (EORTC) recommendations for planning and delivery of high-dose, high precision radiotherapy for lung cancer. Radiother Oncol 2017;124:1–10.
[7] Dehing-Oberije C, De Ruysscher D, van Baardwijk A, Yu S, Rao B, Lambin P. The importance of patient characteristics for the prediction of radiation-induced lung toxicity. Radiother Oncol 2009;91:421–6.
[8] Vogelius IR, Bentzen SM. A literature-based meta-analysis of clinical risk factors for development of radiation induced pneumonitis. Acta Oncol 2012;51:975–83.
[9] Appelt A, Vogelius I, Farr K, Khalil A, Bentzen S. Towards individualized dose constraints; adjusting the QUANTEC radiation pneumonitis model for clinical risk factors. Acta Oncol 2014;53:605–12.
[10] De Ruysscher D, Faivre-Finn C, Nestle U, et al. European Organization for Research and Treatment of Cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. J Clin Oncol 2010;28:5301–10.
[11] Yamamoto T, Kabus S, von Berg J, Lorenz C, Keall PJ. Impact of four-dimensional computed tomography pulmonary ventilation imaging-based functional avoidance for lung cancer radiotherapy. Int J Radiat Oncol Biol Phys 2011;79:279–88.
[12] Defraene G, van Elmpt W, Crijns W, De Ruysscher D. Regional variability in radiation-induced lung damage can be predicted by baseline CT numbers. Radiother Oncol 2017;122:300–6.
[13] Langendijk J, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reductions of side effects : the model-based approach. Radiother Oncol 2013;107:267–73.
[14] Luo Y, El Naqa I, McShan D, Ray D, Lohse I, Matuszak M, et al. Unraveling biophysical interactions of radiation pneumonitis in non-small cell lung cancer via Bayesian network analysis. Radiother Oncol 2017;123:85–92.
[15] Yirmibesoglu E, Higginson D, Fayda M, Rivera M, Halle J, Rosenman J, et al. Challenges scoring radiation pneumonitis in patients irradiated for lung cancer. Lung Cancer 2012;76:350–3.
[16] Christodoulou M, McCloskey P, Stones N, Bayman N, Burt P, Chitallia A, et al. Investigation of a patient-reported outcome tool to assess radiotherapy-related toxicity prospectively in patients with lung cancer. Radiother Oncol 2014;112:244–9.
[17] De Ruysscher D, Dehing C, Yu S, Wanders R, Öllers M, Dingemans A-M, et al. Dyspnea evolution after high-dose radiotherapy in patients with non-small cell lung cancer. Radiother Oncol 2009;91:353–9.
[18] Lievens Y, Pijls-Johannesma M. Health economic controversy and cost-effetciveness of proton therapy. Sem Rad Oncol 2013;23:134–41.
[19] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med 2015;13:1.
[20] Wolthaus J, Schneider C, Sonke JJ, van Herk M, Belderbos J, Rossi M, et al. Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients. Int J Radiat Oncol Biol Phys 2006;65:1560–71.
[21] Verbakel W, van Reij E, Ladenius-Lischer I, Cuijpers J, Slotman B, Senan S. Clinical application of a novel hybrid intensity-modulated radiotherapy technique for stage III lung cancer and dosimetric comparison with four other techniques. Int J Radiat Oncol Biol Phys 2012;83:e297–303.
[22] van Elmpt W, De Ruysscher D, van der Salm A, Lakeman A, van der Stoep J, Emans D, et al. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. Radiother Oncol 2013;104:67–71.
[23] West C, Azria D, Chang-Claude J, Davidson S, Lambin P, Rosenstein B, et al. The REQUITE project : validating predictive models and biomarkers of radiotherapy toxicity to reduce side-effects and improve quality of life in cancer survivors. Clin Oncol 2014;26:739–42.
[24] Bentzen S, Saunders M, Dische S. From CHART to CHARTWEL in non-small cell lung cancer: clinical radiobiological modelling of the expected change in outcome. Clin Oncol 2002;14:372–81.
[25] Matsuo Y, Shibuya K, Nakamura M, Narabayashi M, Sakanaka K, Ueki N, et al. Dose-volume metrics associated with radiation penumonitis after stereotactic body radiation therapy for lung cancer. Int J Radiat Oncol Biol Phys 2012;83:e5445–9.
[26] Wang S, Liao Z, Wei X, Liu H, Tucker S, Hu C, et al. Analysis of clinical and dosimetric factors associated with treatment-related pneumonitis (TRP) in patients with non-small cell lung cancer (NSCLC) treated with concurrent chemotherapy and three-dimensional conformal radiotherapy (3D-CRT). Int J Radiat Oncol Biol Phys 2006;66:1399–407.
[27] Huang EX, Hope AJ, Lindsay PE, et al. Heart irradiation as a risk factor for radiation pneumonitis. Acta Oncol 2011;50:51–60.
[28] Harrell FE. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. Springer; 2001.
[29] Begg CB, Gray R. Calculation of polychotomous logistic regression parameters using individualized regressions. Biometrika 1984;71:11–8.
[30] Hope AJ, Lindsay PE, El Naqa I, et al. Modeling radiation pneumonitis risk with clinical, dosimetric and spatial parameters. Int J Radiat Oncol Biol Phys 2006;65:112–24.
[31] Seppenwoolde Y, De Jaeger K, Boersma LJ, Belderbos JS, Lebesque JV. Regional differences in lung radiosensitivity after radiotherapy for non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2004;60:748–58.
[32] Palma D, Senan S, Tsujino K, Barriger R, Rengan R, Moreno M, et al. Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis. Int J Radiat Oncol Biol Phys 2012;85:444–50.
[33] Rancati T, Ceresoli GL, Gagliardi G, Schipani S, Cattaneo GM. Factors predicting radiation pneumonitis in lung cancer patients: a retrospective study. Radiother Oncol 2003;67:275–83.
[34] Rowell NP, O'Rourke NP. Concurrent chemoradiotherapy in non-small cell lung cancer. The Cochrane Database of Systematic Reviews 2004, Issue 4. Art. No.: CD002140.pub2.
[35] Wijsman R, Dankers F, Troost E, Hoffmann A, van der Heijden E, de Geus-Oei L-F, et al. Inclusion of incidental radiation dose to the cardiac atria and ventricles does not improve the prediction of radiation pneumonitis in advanced stage non-small cell lung cancer patients treated with intensity-modulated radiation therapy. Int J Radiat Oncol Biol Phys 2017;99:434–41.
[36] Defraene G, van Elmpt W, Crijns W, Slagmolen P, De Ruysscher D. CT characteristics allow identification of patient-specific susceptibility for radiation-induced lung damage. Radiother Oncol 2015;117:29–35.
[37] Bernchou U, Schytte T, Bertelsen A, Bentzen SM, Hansen O, Brink C. Time evolution of regional CT density changes in normal lung after IMRT for NSCLC. Radiother Oncol 2013;109:89–94.
[38] Tucker S, Li M, Xu T, Gomez D, Yuan X, Yu J, et al. Incorporating single nucleotide polymorphisms into the Lyman model to improve prediction of radiation pneumonitis. Int J Radiat Oncol Biol Phys 2013;85:251–7.