

Journal of Thoracic Oncology

A competing risk model of first failure site after definitive (chemo) radiation therapy for locally advanced non-small cell lung cancer --Manuscript Draft--

Manuscript Number:	JTO-D-17-01410R1
Full Title:	A competing risk model of first failure site after definitive (chemo) radiation therapy for locally advanced non-small cell lung cancer
Article Type:	Original Article
Keywords:	Locally advanced non-small cell lung cancer; chemo/radiotherapy; competing risk; patient and lesion failure probability; FDG PET
Corresponding Author:	Lotte Nygård, M.D University Hospital Rigshospitalet Copenhagen, DENMARK
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University Hospital Rigshospitalet
Corresponding Author's Secondary Institution:	
First Author:	Lotte Nygård, M.D
First Author Secondary Information:	
Order of Authors:	Lotte Nygård, M.D
	Ivan Vogelius, PhD, DMSci
	Barbara Fischer, M.D, PhD, DMSci
	Andreas Kjær, M.D DMSci
	Seppo Langer, M.D, PhD
	Marianne Aznar, PhD
	Gitte Persson, M.D, PhD
	Søren Bentzen, PhD, DMSci
Order of Authors Secondary Information:	
Manuscript Region of Origin:	DENMARK
Abstract:	<p>Introduction</p> <p>The aim of the study was to build a model of first failure site and lesion specific failure probability after definitive chemo-radiotherapy for inoperable non-small cell lung cancer (NSCLC).</p> <p>Methods</p> <p>We retrospectively analyzed 251 patients receiving definitive chemo/radiotherapy for NSCLC at a single institution between 2009-2015. All patients were FDG PET/CT scanned for radiotherapy planning. Clinical patient data and FDG PET standardized uptake values from primary tumor and nodal lesions were analyzed using multivariate cause-specific Cox regression. In patients experiencing loco-regional failure, multivariable logistic regression was applied to assess risk of each lesion being first site of failure. The two models were used in combination to predict lesion failure probability accounting for competing events.</p> <p>Results</p> <p>Adenocarcinoma had a lower hazard ratio (HR) of loco-regional (LR) failure than squamous cell carcinoma, HR 0.45, 95% CI [0.26; 0.76], $p=0.003$. Distant failures were more common in the adenocarcinoma group, HR 2.21, 95% CI [1.41; 3.48], $p<0.001$. Multivariable logistic regression of individual lesions at the time of first failure showed primary tumors were more likely to fail than lymph nodes, OR 12.8, 95% CI</p>

	<p>[5.10; 32.17], $p < 0.001$. Increasing SUVpeak was significantly associated with lesion failure, OR 1.26 per unit increase, 95% CI [1.12; 1.40], $p < 0.001$. Electronic model: http://bit.ly/LungModelFDG.</p> <p>Conclusions</p> <p>We developed a failure-site specific competing risk model based on patient- and lesion-level characteristics. Failure patterns differed between adenocarcinoma and squamous cell carcinoma, illustrating the limitation of aggregating them into 'non-small-cell lung cancer'. Failure site specific models add complementary information to conventional prognostic models.</p>
--	---

1 **A competing risk model of first failure site after definitive**
2 **(chemo) radiation therapy for locally advanced non-small cell**
3 **lung cancer**
4

5 Lotte Nygård, MD^{a*}, Ivan R. Vogelius, PhD, DMSci^a, Barbara M. Fischer, MD, PhD, DMSci^b, Andreas Kjær, MD,
6 PhD, DMSci^b, Seppo W. Langer, MD, PhD^a, Marianne Aznar, PhD^a, Gitte F. Persson, MD, PhD^a, Søren M. Bentzen,
7 PhD, DMSci^c

8 ^a Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen,
9 Denmark

10 ^b Department of Clinical Physiology, Nuclear Medicine & PET and Cluster for Molecular Imaging, Rigshospitalet,
11 Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark

12 ^c Division of Biostatistics and Bioinformatics, University of Maryland Greenebaum Cancer Center, and Department of
13 Epidemiology and Public Health, University of Maryland School of Medicine, 655 W Baltimore S, MD 21201,
14 Baltimore, USA

15 *Corresponding author: Lotte Nygård, Department of Oncology, Rigshospitalet, Copenhagen University Hospital,
16 Blegdamsvej 9, 2100 Copenhagen, Denmark. E-mail:lotte.nygaard@regionh.dk

17 This work was supported by Danish Cancer Society grant no. R72-A4605-13-S2 and NCI grant no. P30 CA 134274-04.
18 The authors further acknowledge financial support from the Kirsten and Freddy Johansen Foundation. The authors
19 report no conflicts of interests.

20

21

22

23

24

25

26

27

28

29

30

31 **Introduction**

32 The aim of the study was to build a model of first failure site and lesion specific failure probability
33 after definitive chemo-radiotherapy for inoperable non-small cell lung cancer (NSCLC).

34 **Methods**

35 We retrospectively analyzed 251 patients receiving definitive chemo/radiotherapy for NSCLC at a
36 single institution between 2009-2015. All patients were FDG PET/CT scanned for radiotherapy
37 planning. Clinical patient data and FDG PET standardized uptake values from primary tumor and
38 nodal lesions were analyzed using **multivariate** cause-specific Cox regression. In patients
39 experiencing loco-regional failure, **multivariable** logistic regression was applied to assess risk of
40 each lesion being first site of failure. The two models were used in combination to predict lesion
41 failure probability accounting for competing events.

42 **Results**

43 Adenocarcinoma had a lower hazard ratio (HR) of loco-regional (LR) failure than squamous cell
44 carcinoma, HR 0.45, 95% CI [0.26; 0.76], $p=0.003$. Distant failures were more common in the
45 adenocarcinoma group, HR 2.21, 95% CI [1.41; 3.48], $p<0.001$. **Multivariable** logistic regression of
46 individual lesions at the time of first failure showed primary tumors were more likely to fail than
47 lymph nodes, OR 12.8, 95% CI [5.10; 32.17], $p<0.001$. Increasing SUV_{peak} was significantly
48 associated with lesion failure, OR 1.26 per unit increase, 95% CI [1.12; 1.40], $p<0.001$. Electronic
49 model: <http://bit.ly/LungModelFDG>.

50 **Conclusions**

51 We developed a failure-site specific competing risk model based on patient- and lesion-level
52 characteristics. Failure patterns differed between adenocarcinoma and squamous cell carcinoma,
53 illustrating the limitation of aggregating them into 'non-small-cell lung cancer'. Failure site specific
54 models add complementary information to conventional prognostic models.

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

Keywords

Locally advanced non-small cell lung cancer; chemo/radiotherapy; competing risk; patient and lesion failure probability; FDG PET

1. INTRODUCTION

Curative intended chemo-radiotherapy has long been standard of care for inoperable non-small cell lung cancer patients¹. Local or distant progression is frequently seen after therapy, but also death due to competing events is relatively frequent. Despite advances in targeted therapy for a subset of patients with relatively rare genetic mutations, five-year overall survival rates remain unsatisfactory, around 15%^{2,3}. The natural course of disease depends on histology, with adenocarcinoma (AC) metastasizing to the brain more often than squamous cell carcinoma (SCC)^{4,5,6,7}. Nevertheless, adenocarcinoma and squamous cell carcinoma are often lumped together as 'non-small-cell lung cancer' in clinical trials or treatment guidelines.

The Union for International Cancer Control (UICC) TNM classifications^{8,9} provide important prognostic information. However, when deciding on a combination of systemic and local therapies, estimating the risk of loco-regional versus distant recurrence separately would be of obvious clinical interest. Other factors such as tumor volume and number of fluoro-deoxy-glucose/positron emission tomography (FDG-PET) positive lymph nodes have been proposed to improve the pre-treatment prognostic assessment^{10,11}.

Clinical radiotherapy trials have tested dose escalation to the primary lung tumor in an attempt to improve local control and in turn overall survival. However, the large RTOG 0617 randomized trial¹² found no clinical benefit in the dose-escalation arm emphasizing that local intensification may not be a viable strategy for all NSCLC patients. Improved knowledge of the most likely failure sites within a patient may be of relevance for further individualization of treatment options in the future.

97 The aim of the current study was to establish a model of the failure patterns on a patient and lesion
98 level using baseline clinical data and FDG-PET/ computed tomography (CT) scans.

99 2. MATERIALS AND METHODS

100

101 *Patients*

102

103 Data were retrospectively retrieved from medical records and archived scans from consecutive
104 patients diagnosed with inoperable, locally advanced NSCLC and treated at Rigshospitalet,
105 Copenhagen University Hospital from January 2009 to February 2015. In this time period,
106 operability was determined according to clinical stage below IIAN2, co-morbidity of the patient
107 and lung function by a multidisciplinary tumor board consisting of pulmonologists, thoracic
108 surgeons and medical/radiation oncologists. Patients received definitive chemo/radiotherapy or
109 radiotherapy alone. Chemotherapy was either given sequentially prior to radiation or concomitantly
110 with the first cycle of chemotherapy prior to the PET/CT planning scan. Chemotherapy regimens
111 consisted of either cisplatin/vinorelbin or carboplatin/vinorelbin in a three-week schedule, given
112 three to six times. Patients with prior early-stage lung cancer treated with surgery but now
113 candidates for concomitant chemo/radiotherapy (cCRT) were also included in the study. This could
114 for example be due to relapse in a lymph node station or a new primary tumor. These patients were
115 restaged according to the 7th TNM classification from UICC⁸.

116 *FDG-PET*

117 In preparation for radiotherapy planning, an FDG-PET/CT scan was performed on a Siemens
118 Biograph mCT (Siemens Healthineers, Erlangen), on a flat table top in treatment position
119 approximately 60 min after FDG injection (4MBq/kg). An iodine-based contrast medium was

120 injected intravenously during the CT scan according to departmental guidelines. Details on PET
121 and CT data acquisition from our institution is previously published¹³. A maximum of five lesions
122 per patient were evaluated (Two T-sites/three N-sites). If a patient had multiple FDG avid lesions,
123 the five lesions with the highest FDG uptake and the largest diameter were chosen. FDG avid
124 lesions were contoured with region of interest (ROI)s drawn semi-automatically using a threshold
125 of 50 % of the maximum standardized uptake value (SUV_{max}). SUV_{max} , SUV_{peak} , SUV_{mean} and
126 volume (cm^3) were calculated for all individual FDG positive lesions. SUV_{peak} is defined
127 according to the PERCIST criteria¹⁴ as a sphere of 1 cm^3 centering the hottest point in the lesion.
128 The total lesion glycolysis (TLG) is defined as volume x SUV_{mean} . If a primary tumor (T-site) could
129 not be separated from affected lymph nodes (N-site), the lesion was analyzed as a T-site lesion.
130 Data from the PET scan was analyzed on a Mirada XD[®] workstation (version 1.1.0.31).

131 *Radiotherapy*

132 Radiotherapy was given in 2 Gy fractions, five fractions per week to a total of 60-66 Gy. 3D
133 conformal (until January 2014) or volumetric modulated arc therapy (VMAT) planning techniques
134 were applied and based on the midventilation phase of a 4D CT¹⁵. Cone beam CT with tumor match
135 were used for daily image guidance.

136 The study was approved by the Danish Health and Medicines Authority, case no 3-3013-569/1/ and
137 complied with national data protection regulations. According to Danish law, no research ethics
138 approval was necessary due to the retrospective nature of the study.

139 *Statistics and data analysis*

140 Wilcoxon-Mann-Whitney U tests were used for comparison of ordinal or continuous baseline
141 clinical data in AC versus SCC patients. The chi square test was used to test for associations
142 between two categorical variables, Table 1.

143 Date of histology-confirmed diagnosis was used as the start date, and the date of imaging (CT, FDG
144 PET or MR scans) confirming a relapse/progression, was used as failure dates. The time interval
145 between the start date and the earliest radiologically confirmed treatment failure was defined as *time*
146 *to first failure* (TFF). Patients alive with no evidence of disease (NED) were censored at the last
147 follow-up in the clinic. Overall survival (OS) and TFF data was analyzed using Kaplan Meier
148 plots¹⁶ and univariate Cox regression. IBM® SPSS Statistics for Windows, Version 21.0 (Armonk,
149 NY: IBM Corp) was used for these analyses.

150 First site of failure was specified as Tumor-site (T-site) local failure (LF), lymph node (N-site)
151 regional failure (RF), distant metastases, either extra-cranial (ECDM) or intra-cranial (ICDM).
152 Failure within the thorax but outside the radiotherapy planning target volume (PTV) was scored as
153 ECDM. In case of synchronous local and distant failure, the patient was scored as failing distantly
154 (ECDM or ICDM). Patients with loco-regional failure were scored on a lesion level regarding T
155 versus N site lesions and the lesion SUV_{peak} value from the radiotherapy planning scan was
156 calculated for subsequent statistical analysis.

157 Failure sites in adenocarcinoma and squamous cell carcinoma were compared in univariate
158 competing risk modeling. The cuminc function of the CMPRSK package (version 2.2.7) was used
159 to compare the groups with Fine and Gray's test.

160 A statistical plan was made prior to the multivariate cox regression analysis. Coding and variables
161 are listed in Table 2. The variables were preselected to reflect a parsimonious list of clinical
162 prognosticators. Subsequent to data collection changes were made by excluding TLG measures due
163 to strong correlation with GTV and SUV_{peak} yielding unstable models, supplementary Figure 1.
164 SUV_{peak} was chosen prior to seeing the data due to its use in the PERCIST guidelines¹⁴. Also,
165 smoking status in pack-years was excluded as there was no sign of prognostic value (log rank

166 p=0.43 for trend between tertiles) in the current data set. The exclusion of smoking status avoided
167 the complication of imputing six cases with missing data on smoking history.

168 In all cases the model was stratified for the use of concomitant platinum based chemotherapy. In
169 multivariate modeling, ECDM and ICDM were combined to "DM" for power and interpretability of
170 the electronic nomogram. Cause specific Cox multivariate regression was performed using the CSC
171 function of the RiskRegression package, version 1.4.3 in R¹⁷.

172 *Lesion level analysis*

173 Multivariable logistic regression analysis was applied to all lesions in patients with known loco-
174 regional failure. Two variables on lesion specific outcome were included in the analysis: SUV_{peak}
175 and tumor versus node (TvsN) (categorical variable T=1, N=0).

176

$$178 \quad P = \frac{\exp(b_0 + b_1 \times TvsN + b_2 \times SUV_{peak})}{1 + \exp(b_0 + b_1 \times TvsN + b_2 \times SUV_{peak})}$$

177

179 where P is the probability of being first site of failure in the group of patients with known LRF. We
180 interpret P as the conditional failure probability of a lesion, given loco-regional failure. This
181 conditional failure probability, P , is subsequently multiplied by the absolute risk of loco-regional
182 failure from the competing risk analysis to yield individual lesion failure risk at RT planning
183 PET/CT. Thus, the probability of lesion failure, predicted at the time of RT planning, is calculated
184 by multiplying P with the risk of loco-regional failure after 24 months according to the competing
185 risk model.

186 Model calibration was assessed by plotting the lesion predicted probability against observed
187 probability in 8 quantiles.

188 The risk models were published as web applications¹⁸ at <http://bit.ly/LungModel>,
189 <http://bit.ly/LungModelGTV> and <http://bit.ly/LungModelFDG> using the R statistical software
190 package "shiny". We used logistic regression of observed failures vs. predicted risk of lesion failure
191 to estimate the uncertainty of predictions and provide +/- 1.96 times the standard error of the
192 logistic fit as estimate of the uncertainty of the lesion failure probability in the web applications.

193 **RESULTS**

194

195 *Treatment response and follow-up*

196 In this retrospective study 251 patients, of 376 patients screened, were included. The eligibility
197 criteria were: retrievable radiotherapy planning PET/CT scans (5 patients excluded); receiving
198 curative intent radiotherapy (39 patients excluded); available medical records (15 patients
199 excluded); histology of AC or SCC (49 patients excluded), and with FDG PET positive lesions to
200 analyze (17 patients excluded). Patient characteristics are shown in Table 1.

201 Missing data were handled as follows. Performance status (PS) was missing in seven patients but
202 could be retrospectively assessed from electronic file information. Gross tumor volume (GTV) was
203 missing in two patients. A specialist in lung cancer radiology contoured these GTVs retrospectively.
204 Eighteen patients had initial stage IV disease, eleven of which had M1a disease. These lesions
205 received curative intended radiotherapy. Five patients had brain metastases that were either
206 surgically removed or had stereotactic radiotherapy prior to curative intended therapy of the lung
207 lesions. Two patients had M1b disease due to metastases to 1) the adrenal gland and 2) a target in
208 the left breast region. These targets were treated with a stereotactic dose and up to 34 Gy,
209 respectively. Outliers in patient characteristics were age with AC having a larger fraction of

210 younger patients than SCC. T-stage also showed a slight imbalance between the groups,
211 supplementary Figures 2 and 3.

212

213

214 Three patients relapsed in new lymph nodes inside the thorax but outside the PTV. Six patients
215 failed loco-regionally with LN failure outside the thorax (neck, axilla, below the diaphragm). These
216 nine patients were coded with ECDM as first site of failure.

217 Seven patients relapsed inside the thorax but outside the PTV, e.g. contralateral lung (M1a disease)
218 and were coded ECDM. Two of them had other distant metastases (bone and pleura). Seven T-site
219 failures and four N-site failures also failed in distant sites and were scored as distant metastases. See
220 supplementary Table 1 for distribution of first failure sites in AC versus SCC.

221

222 Median time from diagnosis to first relapse was 10.5 months and median time from first relapse to
223 death was 6 months. Overall survival was median 18 months for the whole group of 251 patients
224 with no difference in the two histology groups, HR 0.84, 95% CI [0.62; 1.15]. TFF in the two
225 histology groups were equal, median time 12 months, HR 1.23, 95% CI [0.90; 1.67]. Patients
226 receiving cCRT versus sequential therapy did not have significantly longer OS, HR 1.11, 95% CI
227 [0.81; 1.52].

228 *Competing risk analysis*

229 Competing risk analysis found a strong association between relapse patterns and histology. AC had
230 lower risk of failing LRF compared to SCC, HR 0.45, 95% CI [0.26; 0.76], $p=0.003$. The risk of

231 failing distantly was twice as high in the AC group, HR 2.21, 95% CI [1.41; 3.48], $p < 0.001$. See
232 Table 2.

233 Figure 1 shows the risk of various first-failure types from the competing risk model. The risk of an
234 event increases with time but no major change in risk from two to three years since most failures
235 occur within the first 24 months.

236 Stacked patient level outcome shows by Fine and Grey test, AC to have a higher rate of ICDM,
237 $p = 0.00014$ and ECDM, $p = 0.04$ compared to SCC. SCC tend to fail loco-regionally more often than
238 AC, $p = 0.0002$. There was no difference in Death, NED, $p = 0.18$, see Figure 2.

239 *Lesion risk assessment*

240 Among the 251 patients, 517 lesions were registered with FDG PET uptake from the radiotherapy
241 planning PET/CT scan. Seventy-six lesions (15%) failed in local or regional sites with 60 T-site
242 failures and 16 N-site failures. Logistic regression showed SUVpeak and TvsN to be predictors of
243 lesion failure. T-site lesions were >12 times more likely to fail than N-site lesions, OR 12.8; 95%
244 CI [5.10; 32.17], $p < 0.0001$. Increasing SUVpeak was associated with an increased likelihood of
245 lesion failure, OR 1.26; 95% CI [1.12; 1.40], $p < 0.0001$.

246 We found and analyzed 245 FDG avid lymph nodes, up to three lymph nodes per patient. Of these
247 FDG avid N-site lesions 168 lesions (68.6%) were biopsy proven malignant. Lymph node stations
248 4, 7 and 10 were predominantly represented. Twenty N-site lesions with corresponding positive
249 biopsies relapsed. 238 patients with a total of 493 lesions were included in the overall predicted
250 lesion failure probability analysis. Thirteen patients with no T-site but with 24 nodal lesions were
251 excluded from this analysis. Figure 3 illustrates two patients with different histology and thus
252 difference in lesion failure assessment based on the combined lesion failure probability analysis.

253

254 The lesion model check was performed and showed agreement between calculated risk and actual
255 failure on a lesion level, Figure 4.

256 3. DISCUSSION

257
258 We developed a competing risk model able to estimate the patient level risk of LRF, DM and death
259 NED. Further, the radiation target was subdivided in individual lesions and the model could predict
260 the lesion level risk of failure in the current dataset. Competing risk analysis showed that histology
261 was the strongest predictor for LRF versus DM failure. We found T- versus N-site and SUV_{peak} to
262 be predictive of lesion specific outcome in addition to the patient level prognostics. Internal
263 consistency of the model was examined and we found reasonable agreement between predicted
264 lesion failure probability and actual lesion failure in calibration plots.

265 It is of importance that the model is only used for decision support and should be read with caution.
266 Results from the model are generated under the assumption that all clinical targets, T-and N-sites,
267 are given a homogeneous normo fractionation scheme of 2 Gy times 30-33 times, 5F/W to a total of
268 60 to 66 Gy. It should be stressed that an estimate of low risk of lesion failure does not indicate that
269 it is safe to reduce the dose or otherwise compromise radiation delivery to that lesion.

270 Nevertheless, our results illustrate the importance of differentiating AC and SCC patients since they
271 have different relapse patterns. The aggregate term non-small cell lung cancer may not be helpful in
272 advancing the field.

273 There are numerous reports of the prognostic value of SUV_{max} of the primary lesion and overall cut-
274 off values have been suggested to define good versus poor prognosis^{19,20,21,22}. The lack of external
275 validation should be acknowledged as a limitation of the current study. A number of externally
276 validated prognostic models for overall survival have been published, see for example

277 predictcancer.org. Failure site prognostication as in the current study potentially has, if externally
278 validated, additional clinical interest.

279 We have previously found support for early lesion-specific PET response ($\Delta\text{SUV}_{\text{peak}}$) after one
280 series of chemotherapy to be predictive of failure site in tumor versus nodal sites¹³. A decrease in
281 FDG uptake in both tumor and lymph nodes after radiotherapy was also found to be a prognostic
282 factor for survival and recurrence²³. Looking at subvolumes within the original T-site, other authors
283 have found that high SUV values could identify areas of high local failure risk^{24,25}. This led to
284 studies investigating dose-escalation to these areas hoping to achieve a better local tumor
285 control^{26,27}.

286 When analyzing FDG uptake in lymph nodes²⁸, a risk of false positive findings must be taken into
287 consideration. The rate of false positive lymph node lesions cannot be extracted from our data. SUV
288 measurements varies by glucose blood levels, acquisition modes and reconstruction algorithms of
289 the PET scans which makes data hard to reproduce and define a definitive cut-off value that is
290 prognostic of outcome in a cancer population. An external validation of our findings should be
291 conducted to confirm the rationale in differentiating between histology groups and lesion
292 characteristics and a prospective study could ideally validate our findings of lesion specific risk
293 prediction.

294 Improved estimates of the probability of these competing risks will allow individual treatment
295 approaches that would target the patient's most likely failure type if managed with current standard
296 therapy. The need for such personalized patient selection is particularly evident after the impairment
297 in survival observed for patients in the experimental RT arm of RTOG0617¹², possibly due to
298 excess mortality as a result of the aggressive local treatment intensification. Future randomized

299 trials of local radiotherapy intensification are encouraged to distinguish upon histology and risks of
300 competing events upon inclusion.

301 In conclusion, adenocarcinoma and squamous cell carcinoma of the lung differ in patterns of first
302 failure site after definitive chemo-radiotherapy. Competing risk estimates of DM, LRF and death
303 NED were generated. Moreover, it was possible to estimate risk of failure in sub-lesions of the
304 radiation target based on lesion site, SUV_{peak} and patient level clinical variables.

305

306

307

308

309

310

311

312

313

314

315

316

317

318 REFERENCES

319

- 320 1. Aupérin A et al. Meta-Analysis of Concomitant Versus Sequential Radiochemotherapy in
321 Locally Advanced Non – Small-Cell Lung Cancer. *J Clin Oncol*. 2010;28(13):2181-2190.
322 doi:10.1200/JCO.2009.26.2543.
- 323 2. GLOBOCAN. Lung cancer mortality, worldwide. <http://gco.iarc.fr/today/#>.
- 324 3. Curran WJ, Paulus R, Langer CJ, et al. Sequential vs Concurrent Chemoradiation for Stage
325 III Non-Small Cell Lung Cancer: Randomized Phase III Trial RTOG 9410. *JNCI J Natl*
326 *Cancer Inst*. 2011;103(19):1452-1460. doi:10.1093/jnci/djr325.
- 327 4. Na I II, Lee TH, Choe DH, et al. A diagnostic model to detect silent brain metastases in
328 patients with non-small cell lung cancer. *Eur J Cancer*. 2008;44(16):2411-2417.
329 doi:10.1016/j.ejca.2008.07.003.
- 330 5. Gaspar LE, Chansky K, Albain KS, et al. Time From Treatment to Subsequent Diagnosis of
331 Brain Metastases in Stage III Non – Small-Cell Lung Cancer : A Retrospective Review by
332 the Southwest Oncology Group. *Society*. 2005;23(13):2955-2961.
333 doi:10.1200/JCO.2005.08.026.
- 334 6. Earnest F, Ryu JH, Miller GM, et al. Suspected non-small cell lung cancer: incidence of
335 occult brain and skeletal metastases and effectiveness of imaging for detection--pilot study.
336 *Radiology*. 1999;211(1):137-145. doi:10.1148/radiology.211.1.r99ap34137.
- 337 7. Sánchez de Cos J, Sojo González MA, Montero MV, Pérez Calvo MC, Vicente MJM, Valle
338 MH. Non-small cell lung cancer and silent brain metastasis. Survival and prognostic factors.
339 *Lung Cancer*. 2009;63(1):140-145. doi:10.1016/j.lungcan.2008.04.013.

- 340 8. Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project :
341 Proposals for the Revision of the TNM Stage Groupings in the Forthcoming (Seventh)
342 Edition of the TNM Classification of Malignant Tumours. *J Thorac Oncol*. 2009;2(8):706-
343 714. doi:10.1097/JTO.0b013e31812f3c1a.
- 344 9. Rusch VW, Asamura H, Watanabe H, Giroux DJ, Rami-Porta R, Goldstraw P. The IASLC
345 lung cancer staging project: a proposal for a new international lymph node map in the
346 forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol*.
347 2009;4(5):568-577. doi:10.1097/JTO.0b013e3181a0d82e.
- 348 10. Dehing-Oberije C, De Ruyscher D, van der Weide H, et al. Tumor Volume Combined With
349 Number of Positive Lymph Node Stations Is a More Important Prognostic Factor Than TNM
350 Stage for Survival of Non-Small-Cell Lung Cancer Patients Treated With
351 (Chemo)radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008;70(4):1039-1044.
352 doi:10.1016/j.ijrobp.2007.07.2323.
- 353 11. Dehing-Oberije C, Yu S, De Ruyscher D, et al. Development and external validation of
354 prognostic model for 2-year survival of non-small-cell lung cancer patients treated with
355 chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2009;74(2):355-362.
356 doi:10.1016/j.ijrobp.2008.08.052.
- 357 12. Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal
358 radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without
359 cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a
360 randomised, two-by-two factorial p. *Lancet Oncol*. 2015;16(2):187-199. doi:10.1016/S1470-
361 2045(14)71207-0.
- 362 13. Nygård L, Vogelius IR, Fischer BM, et al. Early lesion-specific 18F-FDG PET response to

363 chemotherapy predicts time to lesion progression in locally advanced non-small cell lung
 364 cancer. *Radiother Oncol*. 2016;118(3):460-464. doi:10.1016/j.radonc.2016.01.009.

365 14. Wahl RL, Jacene H, Kasamon Y, Lodge M a. From RECIST to PERCIST: Evolving
 366 Considerations for PET response criteria in solid tumors. *J Nucl Med*. 2009;50 Suppl 1:122S-
 367 50S. doi:10.2967/jnumed.108.057307.

368 15. Wolthaus JW, Sonke JJ, van Herk M, Belderbos JS, Rossi MM, Lebesque JV DE.
 369 Comparison of different strategies to use four-dimensional computed tomography in
 370 treatment planning for lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2008;70(4):1229-
 371 1238. doi:10.1016/j.ijrobp.2007.11.042.

372 16. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat*
 373 *Assoc*. 1958;53(282):457-481. doi:10.1080/01621459.1958.10501452.

374 17. Team RDC. R: A Language and Environment for Statistical Computing. [http://www.r-](http://www.r-project.org/)
 375 [project.org/](http://www.r-project.org/). 2008. <http://www.r-project.org/>.

376 18. Håkansson K, Rasmussen JH, Rasmussen GB, et al. A failure-type specific risk prediction
 377 tool for selection of head-and-neck cancer patients for experimental treatments. *Oral Oncol*.
 378 2017;74(September):77-82. doi:10.1016/j.oraloncology.2017.09.018.

379 19. Berghmans T, Dusart M, Paesmans M, et al. Primary Tumor Standardized Uptake Value
 380 (SUVmax) Measured on Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) is
 381 of Prognostic Value for Survival in Non-small Cell Lung Cancer (NSCLC): A Systematic
 382 Review and Meta-Analysis (MA) by the European Lung Cancer Working Party for the
 383 IASLC Lung Cancer Staging Project. *J Thorac Oncol*. 2008;3(1):6-12.
 384 doi:10.1097/JTO.0b013e31815e6d6b.

- 385 20. Machtay M, Duan F, Siegel B a, et al. Prediction of survival by [18F]fluorodeoxyglucose
386 positron emission tomography in patients with locally advanced non-small-cell lung cancer
387 undergoing definitive chemoradiation therapy: results of the ACRIN 6668/RTOG 0235 trial.
388 *J Clin Oncol*. 2013;31(30):3823-3830. doi:10.1200/JCO.2012.47.5947.
- 389 21. Vesselle H, Freeman JD, Wiens L, et al. Fluorodeoxyglucose uptake of primary non-small
390 cell lung cancer at positron emission tomography: new contrary data on prognostic role. *Clin*
391 *Cancer Res*. 2007;13(11):3255-3263. doi:10.1158/1078-0432.CCR-06-1128.
- 392 22. Na F, Wang J, Li C, Deng L, Xue J, Lu Y. Primary Tumor Standardized Uptake Value
393 Measured on F18-Fluorodeoxyglucose Positron Emission Tomography Is of Prediction
394 Value for Survival and Local Control in Non-Small-Cell Lung Cancer Receiving
395 Radiotherapy: Meta-Analysis. *J Thorac Oncol*. April 2014:1-9.
396 doi:10.1097/JTO.0000000000000185.
- 397 23. Lopez Guerra JL, Gladish G, Komaki R, Gomez D, Zhuang Y, Liao Z. Large decreases in
398 standardized uptake values after definitive radiation are associated with better survival of
399 patients with locally advanced non-small cell lung cancer. *J Nucl Med*. 2012;53(2):225-233.
400 doi:10.2967/jnumed.111.096305.
- 401 24. Calais J, Thureau S, Dubray B, et al. Areas of High 18F-FDG Uptake on Preradiotherapy
402 PET/CT Identify Preferential Sites of Local Relapse After Chemoradiotherapy for Non-Small
403 Cell Lung Cancer. *J Nucl Med*. 2015;56(2):196-203. doi:10.2967/jnumed.114.144253.
- 404 25. Aerts HJWL, van Baardwijk A a W, Petit SF, et al. Identification of residual metabolic-active
405 areas within individual NSCLC tumours using a pre-radiotherapy (18)Fluorodeoxyglucose-
406 PET-CT scan. *Radiother Oncol*. 2009;91(3):386-392. doi:10.1016/j.radonc.2009.03.006.

- 407 26. Van Elmpt W, De Ruyscher D, Van Der Salm A, et al. The PET-boost randomised phase II
408 dose-escalation trial in non-small cell lung cancer. *Radiother Oncol*. 2012;104(1):67-71.
409 doi:10.1016/j.radonc.2012.03.005.
- 410 27. Partridge M, Ramos M, Sardaro A, Brada M. Dose escalation for non-small cell lung cancer :
411 Analysis and modelling of published literature. *Radiother Oncol*. 2011;99(1):6-11.
412 doi:10.1016/j.radonc.2011.02.014.
- 413 28. Schrevens L, Lorent N, Dooms C, Vansteenkiste J. The role of PET scan in diagnosis,
414 staging, and management of non-small cell lung cancer. *Oncologist*. 2004;9(6):633-643.
415 doi:10.1634/theoncologist.9-6-633.

416

417

418

419

420

421

422

423

424

425

426

427

428 SUPPLEMENTAL DATA

429

430 Supplemental Figure 1.eps

431 Supplemental Figure 2.eps

432 Supplemental Figure 3.eps
433 Supplemental Table 1.word

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

Table 1 Patient and Treatment Characteristics					
	Total (N=251)	%	AC	SCC	P-value
Age at diagnosis (median and range)	66y[37;89]		64y[37;89]	66y[43;89]	0.01
UICC Clinical stage					0.14
I	2	1	1 (0.7%)	1 (0.9%)	
II	28	11	15 (10.4%)	13 (12.1%)	
IIIA/IIIB	203	81	113 (78.5%)	90 (84.1%)	
IV	18	7	15 (10.4%)	3 (2.8%)	
Tumor-stage	Surgery*				0.03
TX	3	1	1 (0.7%)	2 (1.9%)	
T1	(4)	33	13	26 (18.1%)	7 (6.5%)
T2	(2)	53	21	29 (20.1%)	24 (22.4%)
T3	(4)	48	19	32 (22.2%)	16 (15.0 %)
T4		114	45	56 (38.9%)	58 (54.2%)
Nodal-stage					0.38
N0	56	22	29 (20.1%)	27 (25.2%)	
N1	20	8	10 (6.9%)	10 (9.3%)	
N2	121	48	69 (47.9%)	52 (48.6%)	
N3	54	22	36 (25.0%)	18 (16.8%)	
Gender					0.57
Male	152	61	85 (59%)	67 (63%)	
Female	99	39	59 (41%)	40 (37 %)	
Histology					
AC	144	57	144	...	
SCC	107	43	...	107	
WHO performance status					0.26
0	152	60.6	88 (61.1%)	64 (59.8%)	
1	92	36.6	50 (34.7%)	42 (39.3%)	
2	7	2.8	6 (4.2%)	1 (0.9%)	
Smoking pack year (median and range)	40 [0;100]				0.43
(Missing data in 6 patients)					
Chemotherapy regimen					
Cisplatin/vinorelbin	144	57	86 (59.7%)	58 (54.2%)	0.38
Carboplatin/vinorelbin	95	38	53(36.8%)	42 (39.3%)	
No chemotherapy	11	4	5 (3.5%)	6 (5.6%)	
No data available	1	1	0	1	
Concomitant	167	70	97 (69.8%)	70 (70.0%)	0.75

Sequential	72	30	42 (30.2%)	30 (30.0%)
Radiation dose	0.23			
2 Gy x 30 (60 Gy)	9	3	5 (3.5%)	4 (3.7%)
2 Gy x 33 (66 Gy)	240	96	137 (95.1%)	103 (96.3%)
Did not complete (<50 Gy)	2	1	2 (1.4%)	0 (0.0 %)
Target characteristics (Median and range)	AC	SCC	P-value	
GTV	74.9 [3.0; 802.6]	89.0 [3.4; 664.0]	0.27	
PTV	468.9 [50.5; 1605.4]	493.8 [46.0; 1440.4]	0.70	
SUVpeak_T-site #	10.1 [1.0; 58.8]	11.7 [2.6; 43.3]	0.20	
SUVpeak_N-site #	6.9 [0.9; 37.4]	8.5 [1.3; 28.8]	0.38	
MTV	8.3 [0.3; 111.8]	10.4 [0.2; 97.3]	0.44	
TLG_SUM	82.4 [2.0; 825.9]	93.1 [0.96; 1063.4]	0.27	

Table 1. Patient characteristics divided by histology. Categorical variables (Clinical stage, T-stage, N-stage, gender, performance status, chemotherapy regime, and radiation dose) were tested by Chi-square method (p-values in cursive). Non parametric variables (Age, pack year, and target characteristics) were tested by Mann Whitney U test. Significant p-values ($p<0.05$) are highlighted. UICC: The Union for International Cancer Control. AC: Adenocarcinoma. SCC: Squamous cell carcinoma. Gy: Gray. GTV: Gross tumor volume. PTV: Planning target volume. MLD: Mean lung dose. SUV: Standardized uptake value. TLG: Total lesion glycolysis. 13 patients without SUV uptake in T-site (10 VATS) and 3 TX. TX: no primary tumor. Patients with VATS surgery before definitive radiotherapy with no SUVpeak values in their T-sites are listed with their T-stage prior to surgery. # In case of multiple T-site or N-site lesions, the highest value of SUVpeak was chosen.*

Figure 1. Predicted outcome one, two and three years after diagnosis for 251 patients with adenocarcinoma (grey dots) - or squamous cell carcinoma (black dots) of the lung treated at Rigshospitalet from January 2009 to February 2015. Each side of the plot corresponds to the probability of a given endpoint. The black arrow points to the intersection of lines corresponding to 70% probability of no evidence of disease (Alive or dead, NED) – 10% probability of distant metastases (DM), and 20% probability of loco-regional failure (LRF).

Figure 2. Competing risk analysis of first failure site depending on histology. NED: no evidence of disease. ECDM: extra cranial distant metastases. ICDM: intracranial distant metastases. LRF: loco-regional failure.

Figure 3. Illustration of two patients and their lesion failure probabilities.

Adenocarcinoma: Primary tumor in right lung. Predicted failure probability=20% [9 to 31%]

Lymph node metastasis in mediastinum. PFP=2% [0 to 5%]

Squamous cell carcinoma: Primary tumor in left lung. Predicted failure probability=42% [28 to 56%]

Lymph node metastasis in mediastinum. PFP=12% [9 to 15%]

Figure 4. Binominal calibration plot of predicted failure probability on a lesion level and observed lesion failure divided into 8 quantiles including 95% confidence intervals.

Figure 1

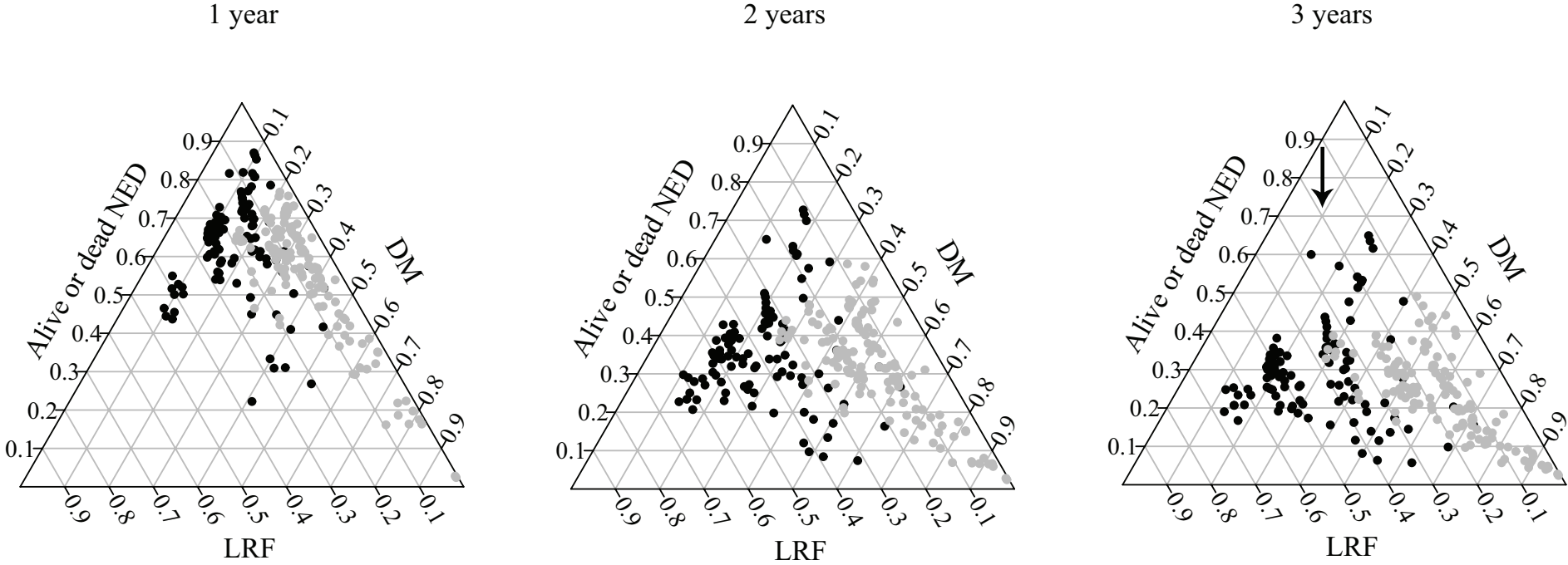


Figure 2

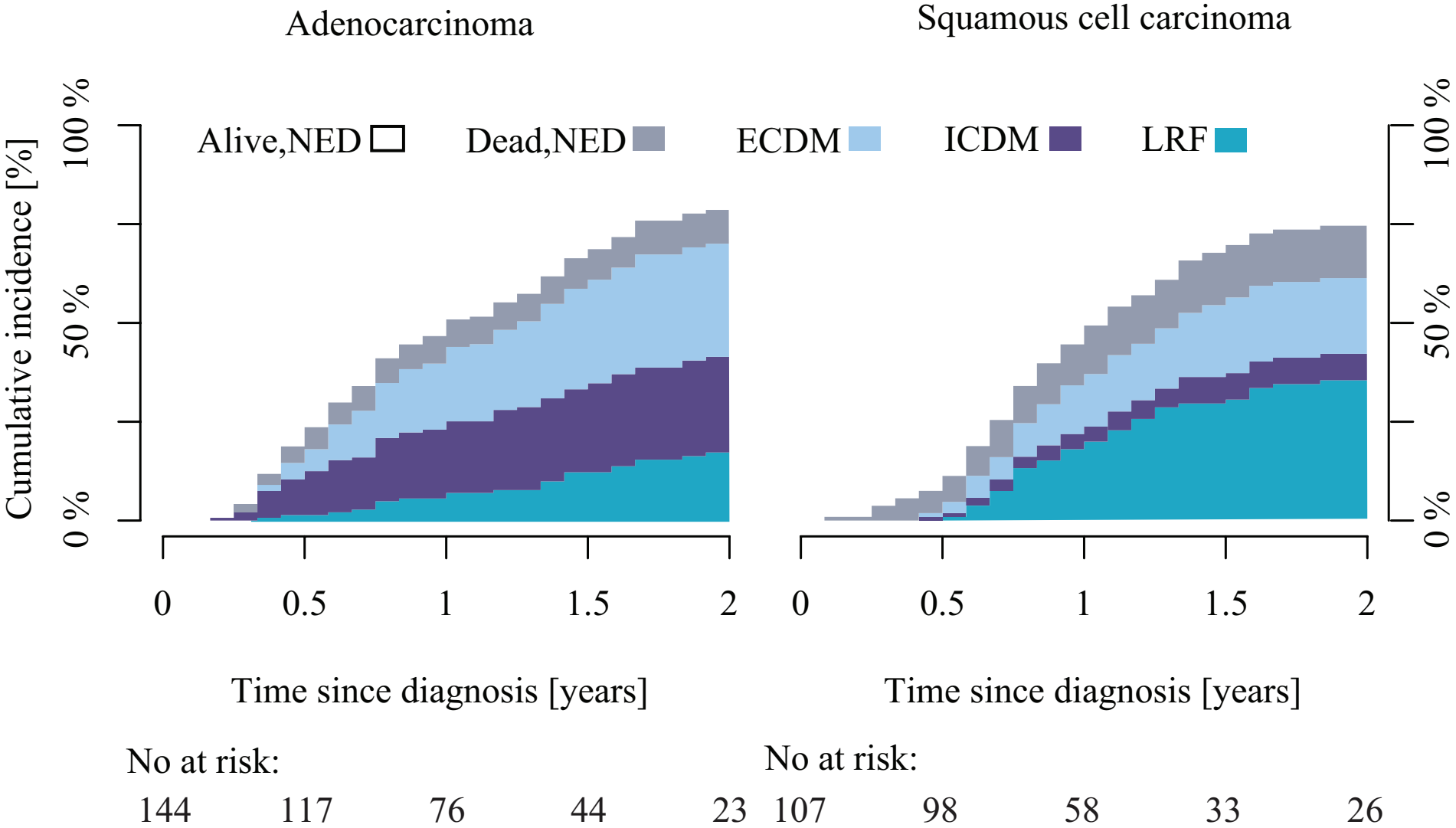


Figure 3

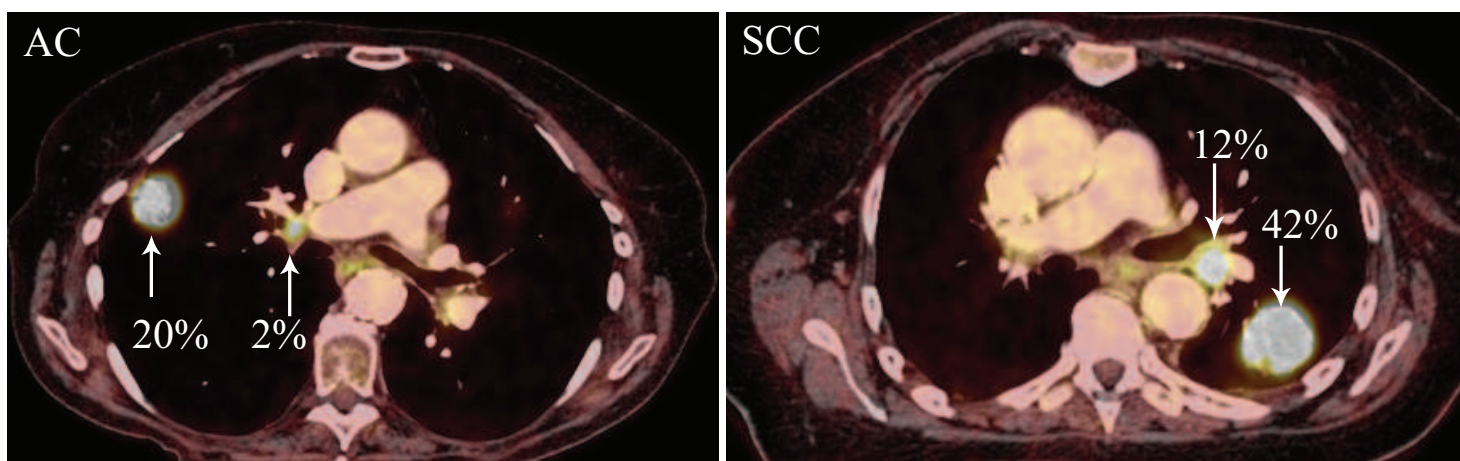


Figure 4

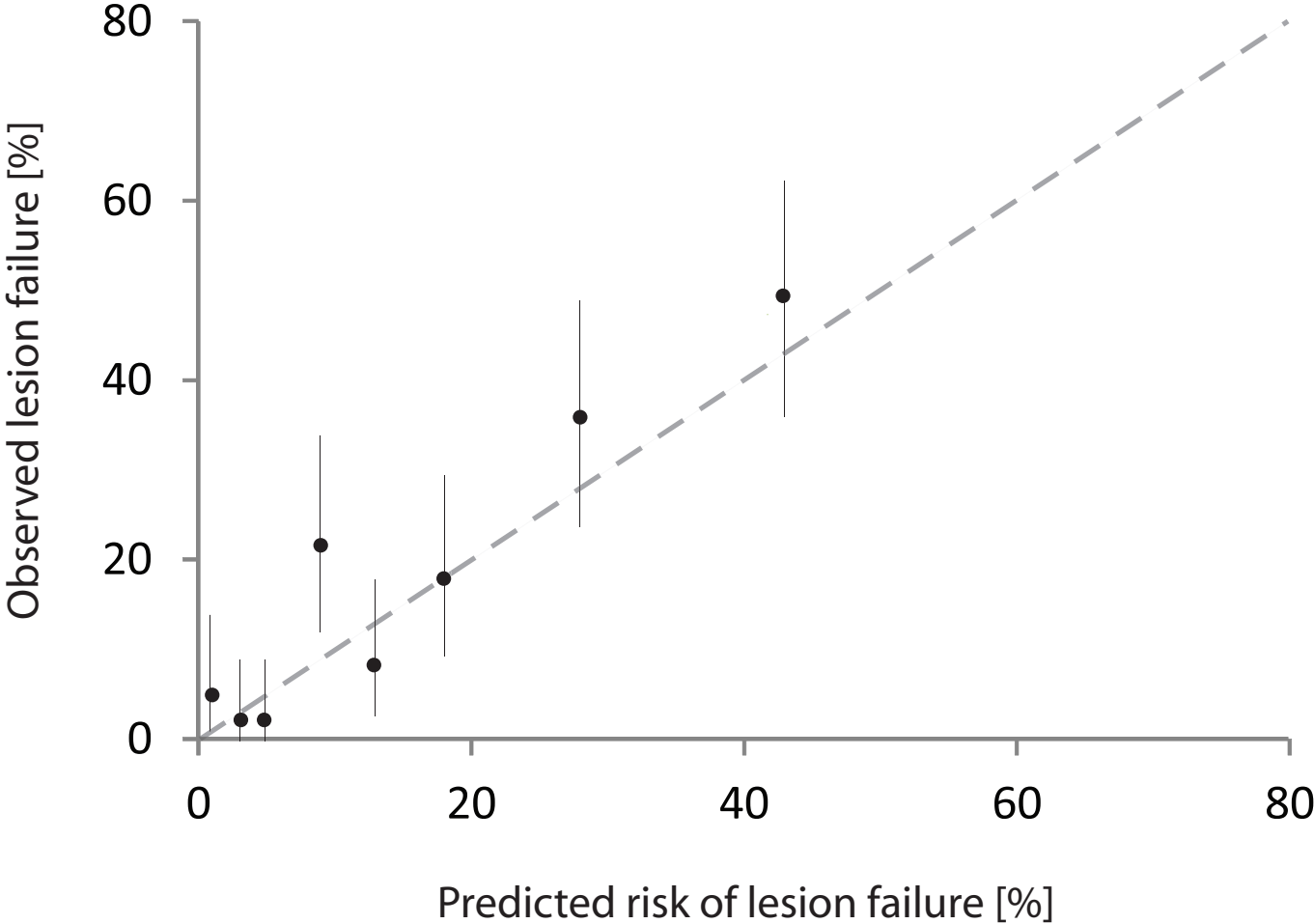


Table 1 Patient and Treatment Characteristics					
	Total (N=251)	%	AC	SCC	P-value
Age at diagnosis (median and range)	66y[37;89]		64y[37;89]	66y[43;89]	0.01
UICC Clinical stage					0.14
I	2	1	1 (0.7%)	1 (0.9%)	
II	28	11	15 (10.4%)	13 (12.1%)	
IIIA/IIIB	203	81	113 (78.5%)	90 (84.1%)	
IV	18	7	15 (10.4%)	3 (2.8%)	
Tumor-stage					0.03
TX	3	1	1 (0.7%)	2 (1.9%)	
T1	(4)	33	13	26 (18.1%)	7 (6.5%)
T2	(2)	53	21	29 (20.1%)	24 (22.4%)
T3	(4)	48	19	32 (22.2%)	16 (15.0 %)
T4		114	45	56 (38.9%)	58 (54.2%)
Nodal-stage					0.38
N0	56	22	29 (20.1%)	27 (25.2%)	
N1	20	8	10 (6.9%)	10 (9.3%)	
N2	121	48	69 (47.9%)	52 (48.6%)	
N3	54	22	36 (25.0%)	18 (16.8%)	
Gender					0.57
Male	152	61	85 (59%)	67 (63%)	
Female	99	39	59 (41%)	40 (37 %)	
Histology					
AC	144	57	144	...	
SCC	107	43	...	107	
WHO performance status					0.26
0	152	60.6	88 (61.1%)	64 (59.8%)	
1	92	36.6	50 (34.7%)	42 (39.3%)	
2	7	2.8	6 (4.2%)	1 (0.9%)	
Smoking pack year (median and range)	40 [0;100]				0.43
(Missing data in 6 patients)					
Chemotherapy regimen					
Cisplatin/vinorelbin	144	57	86 (59.7%)	58 (54.2%)	0.38
Carboplatin/vinorelbin	95	38	53(36.8%)	42 (39.3%)	
No chemotherapy	11	4	5 (3.5%)	6 (5.6%)	
No data available	1	1	0	1	
Concomitant	167	70	97 (69.8%)	70 (70.0%)	0.75

Sequential	72	30	42 (30.2%)	30 (30.0%)
Radiation dose	<i>0.23</i>			
2 Gy x 30 (60 Gy)	9	3	5 (3.5%)	4 (3.7%)
2 Gy x 33 (66 Gy)	240	96	137 (95.1%)	103 (96.3%)
Did not complete (<50 Gy)	2	1	2 (1.4%)	0 (0.0 %)
Target characteristics (Median and range)	AC	SCC		P-value
GTV	74.9 [3.0; 802.6]	89.0 [3.4; 664.0]		<i>0.27</i>
PTV	468.9 [50.5; 1605.4]	493.8 [46.0; 1440.4]		<i>0.70</i>
SUVpeak_T-site [#]	10.1 [1.0; 58.8]	11.7 [2.6; 43.3]		<i>0.20</i>
SUVpeak_N-site [#]	6.9 [0.9; 37.4]	8.5 [1.3; 28.8]		<i>0.38</i>
MTV	8.3 [0.3; 111.8]	10.4 [0.2; 97.3]		<i>0.44</i>
TLG_SUM	82.4 [2.0; 825.9]	93.1 [0.96; 1063.4]		<i>0.27</i>

Table 1. Patient characteristics divided by histology. Categorical variables (Clinical stage, T-stage, N-stage, gender, performance status, chemotherapy regime, and radiation dose) were tested by Chi-square method (p-values in cursive). Non parametric variables (Age, pack year, and target characteristics) were tested by Mann Whitney U test. Significant p-values ($p < 0.05$) are highlighted. UICC: The Union for International Cancer Control. AC: Adenocarcinoma. SCC: Squamous cell carcinoma. Gy: Gray. GTV: Gross tumor volume. PTV: Planning target volume. MLD: Mean lung dose. SUV: Standardized uptake value. TLG: Total lesion glycolysis. 13 patients without SUV uptake in T-site (10 VATS) and 3 TX. TX: no primary tumor. Patients with VATS surgery before definitive radiotherapy with no SUVpeak values in their T-sites are listed with their T-stage prior to surgery. # In case of multiple T-site or N-site lesions, the highest value of SUVpeak was chosen.*

Table 2				
Outcome 1: Loco-regional relapse as first site of failure				
Co-variable	Unit	HR, p; base model	HR, p; Model with GTV	HR, p; Final model with SUVpeak
Age	Year	1.00 [0.96;1.03]	1.00 [0.96;1.03]	0.99 [0.96;1.03]
Histology	AC vs SCC	0.44 [0.26;0.75] p=0.002	0.45 [0.27;0.77] p=0.003	0.45 [0.26;0.76], p=0.003
PS	PS1 and higher vs. PS 0	1.77 [1.06;2.94] p=0.03	1.73 [1.04;2.89] p=0.03	1.73 [1.03;2.92] p=0.04
Clinical Stage	III vs. I and II	3.69 [1.15;11.84] p=0.03	3.61 [1.12;11.59] P=0.03	3.13 [0.97;10.08] p=0.06
	IV vs. I and II	1.9 [0.31;11.60] p=0.49	1.92 [0.31;11.72] p=0.52	1.70 [0.28;10.40] p=0.57
GTV	per 50 cm3 increase	NR	1.04 [0.93;1.15] p=0.52	1.02 [0.91;1.15] p=0.68
SUV peak, T-sites		NR	NR	1.00 [0.96;1.04] p=0.98
Outcome 2: Distant metastases as first site of failure				
Co-variable	Unit	HR, p; base model	HR, p; Model with GTV	HR, p; Model with SUVpeak
Age	Year	0.99 [0.97;1.01]	0.99 [0.96;1.01]	0.99 [0.97;1.02]
Histology	AC vs SCC	2.12 [1.37;3.27] p<0.001	2.43 [1.56;3.78] p<0.001	2.21 [1.41;3.48], p<0.001
PS	PS1 and higher vs. PS 0	0.84 [0.57;1.25] p=0.40	0.77 [0.52;1.14] p=0.20	0.77 [0.51;1.17] p=0.22
Clinical Stage	III vs. I and II	1.56 [0.81;3.03] p=0.20	1.34 [0.69;2.63] P=0.40	1.09 [0.54;2.22] p=0.80
	IV vs. I and II	3.35 [1.47;7.62] p=0.004	3.22 [1.42;7.31] p=0.005	3.19 [1.37;7.44] p=0.007
GTV	per 50 cm3 increase	NR	1.16 [1.07;1.26] p<0.001	1.18 [1.09;1.28] p<0.001
SUV peak, N-sites		NR	NR	1.03 [1.00;1.06] p=0.035
Outcome 3: Dead, NED as first event				
Co-variable	Unit	HR, p; base model	HR, p; Model with GTV	HR, p; Model with SUVpeak
Age	Year	1.02 [0.98;1.07]	1.02 [0.98;1.07]	1.02 [0.98;1.07]
PS	PS1 and higher vs. PS 0	1.59 [0.78;3.24] p=0.21	1.59 [0.78;3.24] p=0.21	1.59 [0.75;3.35] p=0.22
Gender	Female vs Male	1.07 [0.52;2.20] p=0.85	1.07 [0.52;2.20] p=0.85	1.17 [0.55;2.50] p=0.68

Table 2. Sub distribution of hazard ratios in competing risk models. Loco-regional or Distant metastases as first site of failure. Dead, NED as first event. AC: Adenocarcinoma. SCC: Squamous cell carcinoma. HR: hazard ratio, P: P-value, <0.05 is significant. NR: not relevant. GTV: Gross tumor volume. PS: Performance Status. SUVpeak: Standardized Uptake Value. T-sites: Tumor sites. N-sites: Lymph nodes. NED: No evidence of disease.