Validation and comparison of several published prognostic systems for patients with

small cell lung cancer

Paesmans M, Lafitte JJ, Lecomte J, Berghmans T, Efremidis A, Giner V,

Van Cutsem O, Scherpereel A, Meert AP, Leclercq N, Van Houtte P, Sculier JP.

M Paesmans : Data Centre, Institut Jules Bordet, Centre des Tumeurs de l'Université Libre de Bruxelles, Brussels,Belgium
JJ Lafitte : Dept of Pulmonary and Thoracic Oncology, CHRU de Lille University of Lille II, France
J Lecomte : Dept of Pneumology, CHU de Charleroi, Charleroi, Belgium

T. Berghmans, AP Meert, N. Leclercq, JP Sculier : Dept of Intensive Care Unit and Thoracic Oncology, Institut Jules Bordet, Centre des Tumeurs de l'Université Libre de Bruxelles, Brussels, Belgium
A. Efremidis : Medicine Pepartment, Hellenic Cancer Institute – St Savas Oncology Hospital, Athens, Greece
V. Giner : Medical Department, Hospital de Sagunto, Valencia, Spain
O. Van Cutsem : Dept of Pneumology, Clinique Saint-Luc, Bouge, Belgium
A. Scherpereel : Dept of Pulmonary and Thoracic Oncology, CHRU de Lille and University of Lille II, and #INSERM U1019, Centre d'Infection et d'Immunite de Lille, Institut Pasteur de Lille, Lille, France.
P. Van Houtte : Radiotherapy Department, Institut Jules Bordet, Centre des Tumeurs de l'Université Libre de Bruxelles, Brussels, Belgium

Author for correspondance :

Marianne Paesmans

Institut Jules Bordet – Data Centre

Rue Héger-Bordet, 1

B 1000 Brussels Belgium

Tel : +32.2.541.33.99

Fax : +32.2.541.33.97

E-mail : marianne.paesmans@bordet.be

1

Abstract

Objective : To validate and compare published prognostic classifications for predicting survival of patients with small cell lung cancer.

Methods : We pooled data from phase III randomised clinical trials and used Cox models for validation purposes and concordance probability estimates for assessing predictive ability.

Results : We included 693 pts. All the classifications impacted significantly survival with hazard ratios (HR) in the range 1.57-1.68 (all p<0.0001). Median survival times were for the best predicted groups from 16 to 19 months while from 6 to 7 months for the poorest groups. Most of the paired comparisons were also statistically significant. We obtained similar results when restricting the analysis to patients with extensive disease. Multivariate Cox models for fitting survival data were also published. The HRs for a single covariate derived them were : 8.23 (95% CI : 5.88-11.69), 9.46 (6.67-13.50) and, for extensive disease : 5.60 (3.13-9.93), 12.49 (5.57-28.01) and 8.83 (4.66-16.64). Concordance probability estimates ranged from 0.55 to 0.65 (overlapping confidence intervals).

Conclusion : Published classifications are validated and suitable for use at a population level. Expectedly, prediction at an individual level remains problematic. A specific model designed for extensive disease patients does not appear to perform better.

Introduction

Despite a good sensitivity to radiotherapy and chemotherapy, the prognosis of patients with small cell lung cancer is poor with 5 years survival rate being less than 10% [1]. The most reproducible prognostic factor is disease extent which is also the main factor guiding therapy. Very few other prognostic factors have been clearly established : performance status, sex and some routine biological parameters like neutrophil or leucocyte counts or albuminemia [2,3]. Our Group, the European Lung Cancer Working Party (ELCWP), has long been interested in the identification of prognostic factors and about their integration in a classification system that could be used in care for providing information to the patients and in clinical research for stratifying randomization procedures or adjusting treatment comparisons. We published in 2000 [4] such a classification that we constructed using two different statistical strategies : recursive partitioning and amalgamation algorithms (RECPAM) which lead naturally to a patients' classification and Cox regression modelling. The RECPAM technique consists to start with the whole population of patients and to split it according to the prognostic covariate with the highest signifiance level on the outcome (overall survival in the present case). Once the population splitted into two, search for a new split is separately done on the two obtained nodes. The partitioning process goes on iteratively until no further variable is identified as a significant prognostic factor or until the number of patients becomes too small. A second iterative algorithm is then applied as the terminal nodes obtained from different ways in the tree do not necessarily correspond to patients with significantly different survival distributions. After application of the amalgamation algorithm, the classification is immediate. From the data of 763 patients registered in 4 clinical trials

3

(one phase II and 3 phase III), we proposed a classification based on disease extent, Karnofsky performance index, age, sex and relative neutrophil count. Our system includes 4 groups, the one with the best prognosis including only limited disease patients, the one with the worst prognosis including only extensive disease patients and the two intermediate groups being composed of patients of any disease extent. The RECPAM classification as well as the final Cox model were up to now not further tested with the intent to validate them. Moreover, therapeutic standards have nowadays changed with the introduction of combined thoracic radiation and chemotherapy for patients with limited disease. We conducted the present study on a new further series of patients with small cell lung cancer with two objectives. The first one was to validate our published classification as well as our Cox model. The second one was to validate other previously published classifications/models and to compare the different prognostic systems.

Material and methods

We searched in the literature for other published classifications, based on series of patients with a sample size larger than 500 and including only pure prognostic factors (excluding therapeutic covariates). The limit of 500 patients was purely arbitrary but considered as necessary to provide estimates of survival distributions with low variances. We identified 4 other studies with a published classification, all were based on recursive partitioning and amalgamation algorithms [2,5,6,7]. When the publications proposed also a model based on regression analysis [2,4,7], we considered it in our comparisons too.

4

All the selected classifications and models are based on clinical, demographic or routine biological variables. All the RECPAM classifications lead to 4 prognostic groups except the classification proposed by Foster [7] that focuses on patients with extensive disease and has 5 different prognostic levels. The Cox regression models have 4 to 5 factors to explain the distribution of overall survival. The most frequently used covariates are disease extent, sex and age. Other covariates are LDH level, white blood cell or neutrophil counts, creatinine level, alkaline phosphatases or variables linked to disease stage. The full classifications and models are presented in Appendixes 1a and 1b. ~~Tables 1a and 1b show the classifications and models that we aimed to compare in the present study. It should be noted that one classification focused on patients with extensive disease [7].~~

We used as validation series a pooled data base constituted from the data bases of three further clinical trials that the European Lung Cancer Working Party conducted, from 1992 to 2008, in small cell lung cancer. Two of these studies are closed with results published : ELCWP 1923 [8] and ELCWP 1922 [9]. One is still ongoing, ELCWP 1994. The data relative to the patients registered in these studies were not included in the data base used for derivation of our prognostic classification [4]. Those are all phase III trials and their characteristics are described in Appendix 2. Two trials included patients with extensive disease and treatment consisted in chemotherapy alone. The third trial was dedicated to patients with limited disease who were treated by combined chemoradiation and addressed the issue of the possible radiosensitizing effect of cisplatin. For the ongoing trial, we included only in the data base constructed for the present analysis patients randomized before October 31, 2007 in order to get a theoretical follow-up duration greater than 2 years for all patients.

5

Some eligibility criteria were common to the three trials and similar to the eligibility criteria used for the trials considered in [4]. Small cell lung cancer had to be histologically proven and untreated, patients had to have normal haematological, hepatic and renal functions, to have a Karnofsky performance status of at least 60, to have provided informed consent and to be accessible for follow-up. The definition of limited disease was a disease confined to primary site, mediastinum, homolateral subclavicular lymph nodes without malignant pleural effusion in one trial [8]. In the other two studies, it was defined as a disease that could be treated in one radiotherapy field. Criteria of evaluation in the trials included response to treatment after 3 and 6 courses of chemotherapy, progression free survival and overall survival.

Statistical methodology

The present study had, as objectives, using this new series of patients : 1°) the validation of the results the ELCWP published in 2000, 2°) the assessment and comparison of the prognostic values of the selected previously published RECPAM classifications and Cox models. The criterion of evaluation was overall survival, measured since the registration in the trials. All deaths were taken into account.

All the data required for the assessment of the five published classifications were prospectively collected during the conduct of each of the trials.

We used non parametric estimation of the survival distributions and comparison by logrank tests as well as hazard ratios estimates based on Cox regression models. For assessing the prognostic value of the previously published Cox models, we constructed an overall score based on the published regression coefficients. The

6

original scores were transformed to have the same range of theoretical values for allowing comparison of hazard ratios estimates between the models.

As we had one model specific for patients with extensive disease, we tested the general classifications and models both on all the patients and on the subgroup of patients with extensive disease.

In the first two published trials [8,9] that we used for constructing our validation series, we failed to identify any survival benefit between the arms. For the third trial which is ongoing, we did not look at survival comparison between arms. We did not use any stratification by trial for our analyses as one trial was dedicated to limited disease patients only and the other ones exclusively to patients with extensive disease.

To assess the predictive ability of the prognostic covariates, we used the concordance probability estimate [10].

All the significance probabilities that we report are two-tailed, we considered as statistically significant a p value less than 0.05 .

Results

We collected data about a total of 693 patients, 204 from ELCWP 1922 [9], 233 from ELCWP 1923 [8] and 256 from the ongoing trial (ELCWP 1994). There was an intersection between institutions having recruited patients for the derivation series and the institutions contributing to the trials used for the present validation series. However, the overlap was not total and our validation project is intermediate between internal validation and external validation [11]. Patients characteristics are presented in Table 1. Compared to our derivation series, proportion of women increased as well as the rate of patients with a Karnosfky index ≥ 80. The proportion of patients with

7

limited disease was lower than previously. We had some missing data for biological

parameters preventing us to assess the five classifications on all the patients.

Depending on the classifications, rates of missing data ranged from <1% to 12%.

Median length of follow-up was 119 months; death was observed in 646 patients

(93%). Theoretical follow-up was longer than 2 years in 97% of the patients and than

5 years in 85% of the patients.

Distributions of classifications

Table 2 presents the distribution of the classifications on the validation series, for all

patients and for the subgroup of patients with extensive disease. Frequencies for each

category of the classifications are presented after exclusion of patients with missing

data preventing us to allocate them in one of the categories. For each of the

classifications, lower is the category, better is the predicted survival time. Depending

on the classifications, the rates of patients selected as having the best prognosis ranges

from to 17% to 31%, the rate of patients with the worst prognosis varies between 11%

and 26%.

I. Validation study

Figure 1 shows survival curves according to the ELCWP prognostic classification.

Group I had an estimated median survival time of 90 weeks, compared to 48, 34, 28

for groups II, III and IV. The overall comparison is highly significant (p<0.001),

using the prognostic classification as continuous covariate, as well as all the paired

comparisons. Our previous Cox model identified 4 independent prognostic covariates.

Table 3 provides the newly estimated regression coefficients on the validation series if

all 4 covariates are entered together in the model with the same categorization than for

the derivation model. Three of them keep a statistically significant p value (p<0.0001)

8

while the last one, gender, was not significant with a p value of 0.08. The goodness of fit of a model including disease extent alone was improved if the RECPAM classification was used and further improved if the Cox model was used, confirming the usefulness of the integrated models.

II. Prognostic models comparison

In Table 4, we present the estimates of the median survival times in each prognostic category and an overall comparison reflecting the global prognostic value of each RECPAM classification using a non parametric estimation method and the logrank test. When all the patients were analyzed, the overall comparison was in each situation highly significant. The hazard ratio estimates range from 1.57 to 1.87 (change when comparing one category to the adjacent higher one). This means that the hazards for the group with the worst prognosis are 4 to 5 fold higher compared to the group with the best prognosis. No heterogeneity between these hazard ratios was found. As the overall comparisons were all significant, we went for paired comparisons (group i compared to group i+1). The results are presented in Table 5. For the Albain classification [5], there is no statistical evidence that groups II and III have different survival distributions. For the Foster classification [7] focusing on patients with extensive disease, group I cannot be shown different from group II and the comparison between groups II and III is not significant too. Group IV was not analyzed because very few patients in our series belong to that group.

The regression coefficients for the three Cox models were used to calculate an overall prognostic score. The scores were standardized in order to vary on a scale between 0 and 2 (on a patients population between 18 and 85 years). The hazard ratios associated to these scores are shown in Table 6 and are all significantly different from 1. The more recent models [2,7] appear to perform better than the third one [4].

9

Concordance probability estimates

For each classification and model, concordance probability estimates were calculated, predicted survival times were obtained from a Cox model using a single covariate which was a RECPAM classification or a covariate calculated from the regression coefficients of the published Cox models. They are presented in Table 7 with 95% confidence intervals. The Cox models have slightly higher coefficients than the RECPAM classifications. The coefficients are worse when the analysis is restricted to the population of patients with extensive disease, even for the models proposed by Foster that were specifically constructed on patients with extensive disease.

Discussion

~~All of the classifications that we selected proved to be of prognostic value on this new population of patients entered in multicentric clinical studies a well as the companion Cox regression models. For most of them, all the defined groups remained pertinent except those with a low rate of patients like the second group of the IASLC classification or the fourth group of the classification by Foster. That latter one, although specifically designed for extensive disease patients, did not appear to perform better than the general classifications. In each case, the concordance probability estimates show that individual prediction is very difficult and this is particularly true in the population of patients with extensive disease.~~

Prognostic factors studies are numerous in the literature. However, we are lacking from reproducible and well established independent prognostic factors [3]. Indeed, most often, identification of prognostic factors is retrospectively done and there are, few, if any, for patients with small cell lung cancer, phase III prognostic factors studies as labelled by Douglas and Altman [11] that are prospectively conducted with

a priori hypotheses and evaluation of the required sample size according to the hypotheses. Therefore, it is of crucial importance to validate results before making use of them. Further, the independent prognostic value of a factor needs to be validated. In that context, we do not need to validate a single factor but rather a set of prognostic factors and, therefore, a classification or a model. There are several possibilities for validation exercises : internal validation with crossvalidation or bootstrapping techniques or external validation on new series of patients which is the more convincing way to proceed whenever possible [12]. Several levels may be considered for external validation too : using patients recruited at the same institutions and during the same time period is less general than carrying out a validation study on patients from a different origin and developing the disease later than the patients included in the derivation sets. We performed an external validation study of our publication in Cancer in 2000 with an assessment of its historical transportability as defined by Justice [13] with however, institutions having contributed to both series of patients. It should be noted that the classification constructed by the IASLC Lung Cancer Staging Project is not entirely independent as the ELCWP contributed to the constitution of the world wide data base and it is possible that data from the other series were also integrated in that data base. This later classification was already validated as, in the IASLC publication, two thirds of the patients were used for model derivation and one third was kept for validation. Beyond the fact that the validation is not fully external, we should stress that our study is a retrospective one leading to some missing data in the assessment of the classifications/models and to some heterogeneity in the way the required covariates were assessed. Further, our patients population is restricted to a population of patients registered in clinical trials limiting the generalizability of the conclusions.

11

Our validation is successful for both our RECPAM classification and our Cox model in a population of patients included in a clinical trial despite the fact that treatment strategies have evolved with the introduction of combined chemoradiation modalities for limited disease. The prognosis of patients with limited disease has improved and we may hypothesize that this is the joint result of a higher accuracy of staging techniques and a more effective treatment. Further, the four groups of our classification are validated as the ordered paired comparisons are all significant. ~~The intermediate group II is constituted by half of the patients but the 3 other groups are well balanced.~~ Group I, as including only patients with limited disease, might be the target for developing new treatment modalities. A second conclusion for the validation of our models is that the concordance between predicted and true survival times in patients belonging to different risk groups is insufficient for making use of the models at the level of the individual patients. The prognostic covariate built on the Cox model performs slightly better than the RECPAM classification which might be a disappointing result as the RECPAM classification does not seem to benefit from the interaction effects that are naturally integrated during the building process of the classification.

Finally, the performance of our models decreases when we apply it on patients with extensive disease.

Most of these conclusions remain true when we are looking at the other proposed prognostic classifications published before our work or developed recently with a global validation and most the ordered paired comparisons being significant. Among the exceptions is the comparison between groups I and II in the Sagman classification [6]. This may be due to a small proportion of patients in group I and to a lack of power in that setting. The concept of very limited disease may also correspond

nowadays to a highly selected patients population. The other exception is between groups II and III in the IASLC classification [2]; the proportion of patients belonging to group II is also rather small and there might be again an issue of power to confirm the separation between groups II and III. The most relevant difference between groups for [2,4,5], appears to be the identification of the group with the best predicted survival distribution (group I), ie limited disease patients with other favourable features. Depending on the classifications, this group is strictly or less strictly defined with the logical consequences on the size of the group and on its median survival duration.

One of the prognostic classifications [7] we compared was specifically developed for patients with extensive disease with the expectation of being more specific and therefore more accurate. Looking at our results, this is not the case suggesting that other covariates than the one tested should be of importance or that the impact of the other covariates is associated to higher variability. Our sample size also might be insufficient as this classification is based on 5 groups on a restricted validation series. The predictive ability, assessed by the concordance probability estimates of all the models is decreased on the restricted populations of patients with extensive disease.

All the models are based on easy to assess variables (age, sex, routine laboratory parameters) or on variables anyway required to determine tumor stage and therapeutic strategy. It should therefore be recommended to use at least one classification for stratifying patients in a clinical trial or to allow comparisons of patients populations. Indeed, all are of value as there is no relevant difference between the models. To our knowledge, this is the first work validating several classification systems. All of them have been shown to be successful on a further series showing once again that it is less relevant to identify isolated prognostic factors than to integrate them into a prognostic

system and that several prognostic systems although based on different covariates may have comparable discriminant and predictive values. Disease extent, age, sex and performance status are however the cornerstones of the classifications.

None of these classifications takes into account molecular biological factors or gene signatures and there is indeed room for improvement as concordance probability estimates are clearly unsatisfactory. However, as they are very simple to assess, any publication on these new possible prognostic factors should compare the predictive ability of more costly prognostic tools to these simple tests.

Appendix 1a – Prognostic classifications previously published.

Albain (4 groups) [5] – Derivation set n=1137	
Group I	Limited disease, normal LDH level, no pleural effusion, age < 70 years
Group II	All other patients with limited disease
Group III	Extensive disease, normal LDH level
Group IV	Extensive disease, abnormal LDH level
Sagman (4 groups) [6] – Derivation set n=614	
Group I	Very limited disease Or female patients with limited disease, normal LDH level, and WBC < 10 IU
Group II	Male patients with limited disease, normal LDH level and WBC < 10 IU Or limited disease with normal LDH level and WBC ≥ 10 IU Or female patients with extensive disease, ECOG PS 0 or 1, no liver metastasis
Group III	Limited disease, elevated LDH level Or female patients with extensive disease, ECOG PS 0 or 1, liver metastasis Or male patients with extensive disease, ECOG PS 0 or 1 Or extensive disease, ECOG PS 2 or 3, normal alkaline phosphatases level
Group IV	Extensive disease, ECOG PS 2 or 3 and elevated alkaline phosphatases level
Paesmans (4 groups) [4] – Derivation set n=763	
Group I	Limited disease, KPS ≥ 80, age < 60 years
Group II	Limited disease, KPS ≥ 80, age ≥ 60 years Or limited disease, KPS ≤ 70 Or extensive disease, KPS ≥ 80, neutro ≤ 75%
Group III	Extensive disease, KPS ≥ 80, neutro > 75% Or extensive disease, KPS ≤ 70, female
Group IV	Extensive disease, KPS ≤ 70, male
Sculier (4 groups) [2] – Derivation set n=4359	
Group I	Limited disease, ECOG PS 0 Or limited disease, ECOG PS 1-2, age < 65 years
Group II	Limited disease, ECOG PS 1-2, age ≥ 65 years Or extensive disease, ECOG PS < 1, female, age < 65 years
Group III	Extensive disease, ECOG PS 0, female, age ≥ 65 years Extensive disease, ECOG PS 0, male Extensive disease, ECOG PS 1, age < 70 years
Group IV	Limited disease, PS 3-4 Extensive disease, ECOG PS 1, age ≥ 70 years Extensive disease, ECOG PS ≥ 2
Foster (5 groups) - ED SCLC only [7] – Derivation set n=910	
Group I	<2 metastatic sites and creatinine < upper normal limit
Group II	Female patients with ≥ 2 metastatic sites
Group III	Male patients with ≥ 2 metastatic sites, ECOG PS 0 or 1
Group IV	< 2 metastatic sites, creatinine > upper normal limit
Group V	Male patients with ≥ 2 metastatic sites and ECOG PS 2

The second column of the table describes the characteristics of the patients included in the group labelled in the first column for each of the classifications. In all

classifications, group I has the best predicted prognosis and the last group has the worst predicted prognosis.

Abbreviations :

WBC : white blood cell count   Neutro : neutrophil count

PS : performance status   KPS : Karnofsky performance status

ED : extensive disease   SCLC : small cell lung cancer

LDH : lactate dehydrogenase

Very limited disease : defined by the authors as limited disease without mediastinal involvement or pleural effusion or ipsilateral supraclavicular adenopathy

Appendix 1b – Cox models previously published.

Covariate	Regression coefficient
Paesmans [4]	
Karnofsky PS < 80	0.33
Male gender	0.34
Extensive disease	0.41
Neutrophil rate > 75%	0.22
Sculier [2]	
Age (1 year increase)	0.010
Male gender	0.22
Extensive disease	0.76
PS 1	0.31
PS2	0.66
PS 3-4	1.24
Foster [7]	
Age (1 year increase)	0.014
Male	0.24
PS 1	0.17
PS 2	0.48
Creatinine level above UNL	0.29
≥ 2 metastatic sites	0.24

Are reported in the table the regression coefficients for each covariate included in a multivariate prognostic model. These regression coefficients may be used to construct one single prognostic variables.

Albain and Sagman reported models specific for LD and ED. As our purpose was to validate the prognostic classifications and to compare their predictive ability with the one of Cox models based on the same patients populations, we did not consider them.

Appendix 2 – Characteristics of the clinical trials

Trial	Patients population	Therapeutic regimen	Registration period	N patients
1922 [9]	Limited disease	Cisplatin 90 mg/m2 d1 Etoposide 100 mg/m2 d1-d3 Vs Cisplatin 6 mg/m2 d1-5, d 8-12, d15-19 Etoposide 100 mg/m2 d1-3 (first course) Chest RT 39.90 Gy 15 fractions Standard chemotherapy in both arms after 1st course	March 1993 to March 2006	204
1923 [8]	Extensive disease	Epirubicin 90 mg/m2 Vindesine 3 mg/m2 Ifosfamide 5 g/m2 every 3 weeks vs every 2 weeks with GM-CSF support vs every 2 weeks with oral antibiotics	April 1993 to April 2000	233

		support		
1994	Extensive disease	Cisplatin 90 mg/m2 d1 Etoposide 100 mg/m2 d1-3 Vs Ifosfamide 1.5 g/m2 d1-3 Etoposide 100 mg/m2 d1-3 Epirubicin 60 mg/m2 d1	September 1999 - ongoing	256

Table 1 - Patients characteristics

Trial reference Characteristic	1922 (n=204)	1923 (n=233)	1994 (n=256)	Total (n=693)	[4] (n=763)	[2] (n=6609)
Gender						
Male	163 (80%)	196 (84%)	211 (82%)	570 (82%)	689 (90%)	4368 (66%)
Female	41 (20%)	37 (16%)	45 (18%)	123 (18%)	74 (10%)	2241 (34%)
Age (years)						
Median	59	61	61	60	NR	NR
Range	32-75	34-74	40-83	32-84		
Karnofsky performance status						
60-70	14 (7%)	57 (24%)	64 (25%)	135 (19%)	250 (33%)	1409 (21%)
80-90	190 (93%)	176 (76%)	192 (75%)	558 (81%)	513 (67%)	5200 (79%)
100	88 (43%)	41 (18%)	31 (12%)	160 (23%)	NR	2039 (31%)
Disease extent						
Limited disease	204 (100%)	18 (8%)	16 (6%)	238 (34%)	365 (48%)	2870 (43%)
Very limited disease	204	1	0	205	NR	NR
Extensive disease	-	215 (92%)	240 (94%)	455 (66%)	398 (52%)	3739 (57%)
> 2 metastatic sites		30	53	83	NR	NR
Pleural effusion						
Not documented	204 (100%)	200 (86%)	219 (86%)	623 (90%)	NR	NR
Yes	-	33 (14%)	37 (14%)	70 (10%)		
White blood cell count						
< 10 IU	152 (75%)	155 (67%)	153 (60%)	460 (66%)	516	NR
≥ 10 IU	52 (25%)	78 (33%)	100 (39%)	230 (33%)	241	
Not reported	-	-	3 (1%)	3 (<1%)	6	
Relative neutrophil count						
≤ 75%	166 (81%)	154 (66%)	152 (59%)	472 (68%)	528 (69%)	NR
> 75%	35 (17%)	74 (32%)	68 (27%)	177 (26%)	212 (28%)	
not assessed	3 (1%)	5 (2%)	36 (14%)	44 (6%)	23 (3%)	
Alkaline phosphatases						
Normal	101 (50%)	69 (30%)	107 (42%)	277 (40%)	297	NR
Abnormal	77 (38%)	158 (68%)	144 (56%)	379 (55%)	431	
Not reported	26 (13%)	6 (3%)	5 (2%)	37 (5%)	35	
Creatinine						
Normal	199 (98%)	225 (97%)	249 (97%)	673 (97%)	NR	NR
Abnormal	3 (1%)	6 (3%)	4 (2%)	13 (2%)		
Not reported	2 (1%)	2 (1%)	3 (1%)	7 (1%)		
LDH						
Normal	144 (71%)	123 (53%)	127 (50%)	394 (57%)	NR	NR
Abnormal	30 (15%)	88 (38%)	98 (38%)	216 (31%)		
Not reported	30 (15%)	22 (9%)	31 (12%)	83 (12%)		

NR=not reported

19

Table 2 – Distributions of the classifications

Classification	Group I	Group II	Group III	Group IV	Group V	Missing
All patients (n=693)						
Albain [5]	145 (21%)	60 (9%)	227 (33%)	178 (26%)	-	83
Sagman [6]	212 (31%)	37 (5%)	364 (53%)	73 (11%)	-	7
Paesmans [4]	113 (17%)	352 (53%)	103 (15%)	98 (15%)	-	27
Sculier [2]	195 (28%)	53 (8%)	279 (40%)	166 (24%)	-	-
Patients with extensive disease (n=455)						
Albain [5]	-	-	227 (56%)	178 (44%)	-	50
Sagman [6]	-	23 (5%)	356 (79%)	73 (16%)	-	3
Paesmans [4]		227 (53%)	103 (24%)	98 (23%)	-	27
Sculier [2]		10 (2%)	279 (61%)	166 (37%)	-	-
Foster [7]	232 (51%)	34 (7%)	131 (29%)	4 (1%)	52 (12%)	2

Table 3 – Prognostic value of the 4 prognostic independent covariates from

(Paesmans 2000)

Characteristic	HR, 95% CI and p value Derivation series n=763, 45 censored observations	HR, 95% CI and p value Validation series n=693, 47 censored observations
Karnofsky PS $\geq$ 80	0.73 (0.60-0.86), p<0.001	0.56 (0.46-0.70), p<0.0001
Female gender	0.71 (0.54-0.94), p=0.02	0.83 (0.66-1.02), p=0.08
Limited disease	0.66 (0.57-0.78), p<0.001	0.45 (0.38-0.51), p<0.0001
Neutrophil rate $\leq$ 75%	0.80 (0.67-0.96), p=0.05	0.68 (0.57-0.82), p<0.0001

Table 4 – Estimates of median survival times in each category and overall comparison

Classification	I	II	III	IV	V	p	HR
All patients							
Albain [5]	17.0 (14.4-20.8)	10.6 (7.5-12.5)	10.5 (9.4-11.5)	7.3 (6.4-11.5)	-	<0.0001	1.58 (1.45-1.72)
Sagman [6]	15.7 (13.4-19.1)	13.3 (10.6-16.7)	9.5 (8.8-10.0)	6.4 (4.7-7.1)	-	<0.0001	1.60 (1.47-1.75)
Paesmans [4]	19.4 (13.9-23.9)	11.1 (10.6-11.8)	7.1 (6.6-8.7)	6.4 (4.0-7.3)	-	<0.0001	1.68 (1.53-1.84)

Sculier [2]	15.9 (13.5-19.9)	10.4 (7.6-13.0)	10.5 (9.6-11.2)	6.8 (6.1-7.8)	-	<0.0001	1.57 (1.45-1.71)
Patients with extensive disease							
Albain [5]	-	-	10.5 (9.4-11.5)	7.3 (6.4-8.3)	-	<0.001	1.87 (1.51-2.31)
Sagman [6]	-	15.9 (11.0-21.3)	9.6 (9.0-10.3)	6.4 (4.7-7.1)	-	<0.001	1.70 (1.35-2.13)
Paesmans [4]	-	10.9 (10.0-11.6)	7.1 (6.6-8.7)	6.3 (4.0-7.3)	-	<0.001	1.44 (1.27-1.64)
Sculier [2]	-	-	10.5 (9.6-11.2)	6.8 (6.1-7.8)	-	<0.001	1.72 (1.41-2.10)
Foster [7]	9.7 (8.8-10.7)	9.5 (7.0-11.1)	9.4 (8.6-10.4)	-	6.4 (3.3-7.8)	<0.001	1.23* (1.12-1.35)

For the calculation of the overall HR, only patients with all classifications available were considered (n=588 for all patients, n=383 for patients with extensive disease). The calculation of the HR was made using the classification as a "continuous" covariate.

*Groups IV and V pooled together.

Table 5 – Paired comparisons

Classification	I versus II	II versus III	III versus IV	IV versus V
All patients				
Albain [5]	1.93 (1.39-2.64)	1.21 (0.90-1.66)	1.88 (1.53-2.31)	-
Sagman [6]	1.47 (1.00-2.11)	1.73 (1.23-2.51)	1.67 (1.29-2.15)	-
Paesmans [4]	2.24 (1.77-2.87)	1.74 (1.39-2.18)	1.33 (1.00-1.77)	-
Sculier [2]	2.17 (1.57-2.96)	1.13 (0.84-1.53)	1.77 (1.45-2.15)	-
Patients with extensive disease				
Albain [5]	-	-	1.88 (1.53-2.31)	-
Sagman [6]	-	-	1.70 (1.30-2.18)	-
Paesmans [4]	-	1.52 (1.20-1.92)	1.33 (1.00-1.77)	-
Sculier [2]		-	1.77 (1.45-2.15)	-
Foster [7]	1.18 (0.79-1.69)	1.04 (0.71-1.57)	-	1.59 (1.15-2.17)

For each comparison, the HR with its 95% CI is reported.

Table 6 – Predictive value of the previously published Cox models

| Classification	HR	95% CI	p
All patients			
Paesmans	8.23	5.88-11.69	<0.001
Sculier	9.46	6.67-13.50	<0.001
Patients with extensive disease			
Paesmans [4]	5.60	3.13-9.93	<0.001
Sculier [2]	12.49	5.57-28.01	<0.001
Foster [7]	8.83	4.66-16.74	<0.001

HR : HR for a "continuous" covariate calculated as linear combination of the independent prognostic covariates weitghted by their regression coefficients.

Table 7 – Concordance probability estimates

RECPAM classifications	CPE	95% CI
All patients		
Albain [5]	0.63	0.61-0.65
Sagman [6]	0.62	0.61-0.65
Paesmans [4]	0.62	0.60-0.64
Sculier [2]	0.63	0.61-0.65
Patients with extensive disease		
Albain [5]	0.58	0.55-0.60
Sagman [6]	0.55	0.53-0.57
Paesmans [4]	0.58	0.55-0.60
Sculier [2]	0.57	0.54-0.59
Foster [7]	0.56	0.54-0.59
Cox models	CPE	95% CI
All patients		
Paesmans [4]	0.64	0.62-0.66
Sculier [2]	0.65	0.63-0.67
Patients with extensive disease		
Paesmans [4]	0.58	0.55-0.60
Sculier [2]	0.59	0.56-0.62
Foster [7]	0.60	0.57-0.63

The concordance probability estimate represents the proportion of patients where the predicted survival times and the observed survival times are ordered the same way.

Figure 1 – Survival curves according to the ELCWP prognostic classification
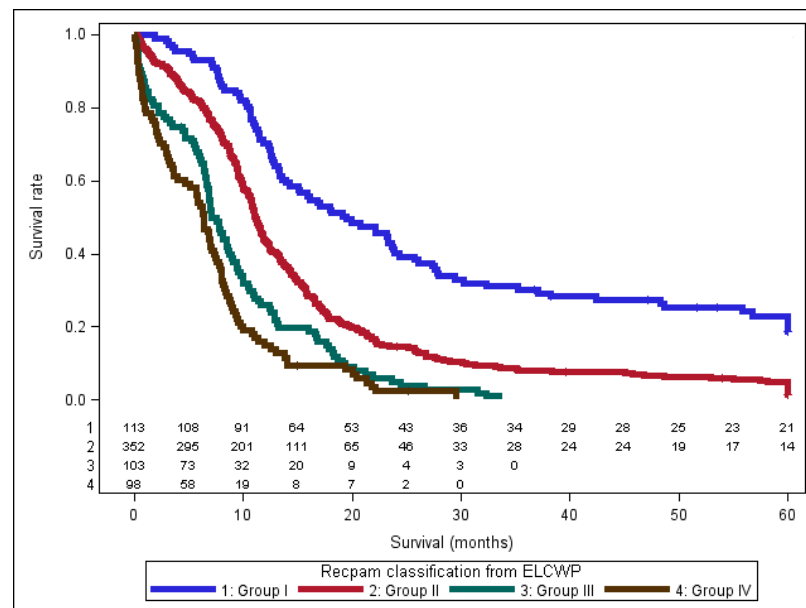
Bibliography

1. Shepherd FA,Crowley J, Van Houtte P, Postmus PE, Carney D, Chansky K, Shaikh Z, Goldstraw P; International Association for the Study of Lung Cancer International Staging Committee and Participating Institutions. The International Association for the Study of Lung Cancer lung cancer staging project: proposals regarding the clinical staging of small cell lung cancer in the forthcoming (seventh) edition of the tumor, node, metastasis classification for lung cancer. J Thorac Oncol 2007; 2(12):1067-77.

2. Sculier JP, Chansky K, Crowley JJ, Van Meerbeeck J, Goldstraw P; International Association for the Study of Lung Cancer International Staging Committee and Participating Institutions. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th Edition of the TNM Classification of Malignant Tumors and the proposals for the 7th Edition. J Thorac Oncol 2008; 3(5):457-66.

23

3. Paesmans M. Prognostic factors in lung cancer. Rev Mal Respir. 2005;22(6):8S76-80.

4. Paesmans M, Sculier JP, Lecomte J, Thiriaux J, Libert P, Sergysels R, Bureau G, Dabouis G, Van Cutsem O, Mommen P, Ninane V, Klastersky J.Prognostic factors for patients with small cell lung carcinoma : analysis of a series of 763 patients included in 4 consecutive prospective trials with a minimum follow-up of 5 years. Cancer 2000; 89(3):523-33.

5. Albain KS, Crowley JJ, Leblanc M, Livingston RB. Determinants of improved outcome in small-cell lung cancer: an analysis of the 2,580-patient Southwest Oncology Group data base. J Clin Oncol 1990; 8(9):1563-74.

6. Sagman U, Maki E, Evans WK, Warr D, Shepherd FA, Sculier JP, Haddad R, Payne D, Pringle JF, Yeoh JL, Ciampi A, Deboer G, McKinney S, Ginsberg R, Feld R. Small-cell carcinoma of the lung: derivation of a prognostic staging system. J Clin Oncol 1991; 9(9):1639-49.

7. Foster NR, Mandrekar SJ, Schild SE, Nelson GD, Rowland KM, Deming RL, Kozelsky TF, Marks RS, Jett JR, Adjei AA. Prognostic factors differ by tumor stage for small cell lung cancer: a pooled analysis of North Central Cancer Treatment Group trials. Cancer 2009;115(12):2721-31.

8. Sculier JP, Paesmans M, Lecomte J, Van Cutsem O, Lafitte JJ, Berghmans T, Koumakis G, Florin MC, Thiriaux J, Michel J, Giner V, Berchier MC, Mommen P, Ninane V, Klastersky J; European Lung Cancer Working Party. A three-arm phase III randomised trial assessing, in patients with extensive-disease small-cell lung cancer, accelerated chemotherapy with support of haematological growth factor or oral antibiotics. Br J Cancer 2001; 85(10):1444-51.

9. Sculier JP, Lafitte JJ, Efremidis A, Florin MC, Lecomte J, Berchier MC, Richez M, Berghmans T, Scherpereel A, Meert AP, Koumakis G, Leclercq N, Paesmans M, Van Houtte P ; European Lung Cancer Working Party (ELCWP). A phase III randomised study of concomitant induction radiochemotherapy testing two modalities of radiosensitisation by cisplatin (standard versus daily) for limited small-cell lung cancer. Ann Oncol 2008; 19(10):1691-7.

10. Harrel FE. Tutorial in biostatistics : multivariable prognostic models : issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996; 15:361-87.

11. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. Br J Cancer 1994; 69(6):979-85.

12. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research : validating a prognostic model. BMJ 2009;338:1432-5.

13. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999;130(6):515-24.