

Prognosis in lung cancer: physicians' opinions compared with outcome and a predictive model

M F Muers, P Shevlin, J Brown on behalf of the participating members of the Thoracic Group of the Yorkshire Cancer Organisation

Abstract

Background – Although the study of prognostic factors in small cell lung cancer has reached the stage where they are used to guide treatment, fewer data are available for non-small cell lung cancer. Although correct management decisions in non-small cell lung cancer depend upon a prognostic assessment by the supervising doctor, there has never been any measurement of the accuracy of physicians' assessments.

Methods – A group of consecutive patients with non-small cell lung cancer was studied and the predictions of their physicians as to how long they would survive (in months) was compared with their actual survival. A prognostic index was also developed using features recorded at the patients' initial presentation.

Results – Two hundred and seven consecutive patients diagnosed and managed as non-small cell lung cancer, who did not receive curative treatment for their condition, were studied. Of the 196 patients whose date of death was known, physicians correctly predicted, to within one month, the survival of only 19 patients (10%). However, almost 59% of patients (115/196) had their survival predicted to within three months and 71% (139/196) to within four months of their actual survival. Using Cox's regression model, the sex of the patient, the activity score, the presence of malaise, hoarseness and distant metastases at presentation, and lymphocyte count, serum albumin, sodium and alkaline phosphatase levels were all identified as useful prognostic factors. Three groups of patients, distinct in terms of their survival, were identified by the use of these items. When the prediction of survival made by the physician was included as a prognostic factor in the original model, it was shown to differentiate further between the group with a poor prognosis and the other two groups in terms of survival.

Conclusions – Physicians were highly specific in identifying patients who would live less than three months. However, they had a tendency to overestimate survival in these patients, failing to identify almost half the patients who actually died within this time. Both the physicians and the prognostic factor model gave similar performances in that they were more successful in identifying patients who had a short

time to survive than those who had a moderate or good prognosis. Physicians appear to use information not identified in the prognostic factor analysis to reach their conclusions.

(Thorax 1996;51:894-902)

Keywords: non-small cell lung cancer, prognostic factors.

Although lung cancer rates, particularly in those aged under 60, are beginning to decline,¹ it remains the commonest cancer with about 35 000 deaths per year in Britain. Under 20% of tumours are resectable, and only 25-35% of operated patients are cured.²⁻³ The prognosis of the rest is poor with a five year survival of 2%.⁴ The most common management problem is therefore not to attempt to prolong survival, but to relieve symptoms and maintain as good a quality of life for as long as possible.

An assessment of prognosis, either explicit or implicit, underpins all these management decisions and influences the choice of treatment, its timing, and duration. An accurate prognosis is also important to guide discussion with the patient. This is a much larger problem than that of staging lung cancer preoperatively which has been analysed in detail elsewhere.⁵⁻⁷

In recent years there has been much study of the clinical features at presentation in small cell lung cancer which have the largest impact on survival – that is, prognostic factors.⁸ Much less work has been done in non-small cell lung cancer since the early report of Stanley⁹ whose study population consisted of patients with non-small cell and small cell lung cancer. There is, however, no reason to suppose that a prognostic factor index cannot be developed in non-small cell lung cancer.

The clinical usefulness of a predictive model depends on it being superior in practice to a physician's opinion. To our knowledge there has never been an assessment of a physician's ability to estimate survival in lung cancer. Thus, in this study, physicians' estimates of survival made at the time of treatment decision were compared with the actual survival of the patient and also with a predictive index based on optimal grouping of prognostic factor measurements made at presentation.

Methods

Sequential patients were recruited from the clinics of four physicians (two district general

Thoracic Group of the
Yorkshire Cancer
Organisation
M F Muers
P Shevlin
J Brown

Correspondence to:
Dr M F Muers, Respiratory
Unit, Regional
Cardiothoracic Centre,
Killingbeck Hospital, Leeds
LS14, 6UQ, UK.

Received 27 September 1995
Returned to authors
15 December 1995
Revised version received
25 March 1996
Accepted for publication
2 April 1996

hospitals and two teaching hospitals) and two radiotherapists/oncologists (both with clinics at a district general hospital and a regional radiotherapy centre). In total, nine doctors examined and recorded their prediction of survival for at least one patient.

The criterion for initial entry to the study was that the patient was being managed as a case of non-small cell lung cancer. For the purpose of this examination the criteria were further extended so that only patients who did not receive curative treatment were included in the study population.

At the end of the study the presenting features and the clinical course of all patients without histological confirmation were reviewed. Patients were only finally accepted as having non-small cell lung cancer without histological confirmation if (1) they had no evidence of any other disease to explain their clinical presentation and course; (2) they had an abnormal chest radiograph on presentation which worsened during the period of observation; (3) they had a clinical course during observation which was compatible with progressive carcinoma or they died during the period of observation with evidence of an enlarging tumour or metastatic disease; (4) their neuron specific enolase (NSE) value was not more than 12 ng/ml (values above this in our laboratory are strongly associated with neuroendocrine tumours such as small cell lung cancer¹⁰); (5) they had no evidence of a primary tumour elsewhere; and (6) the clinical course of the patients was indistinguishable from those with histological confirmation.

Patient data were recorded at the time of the management decision on a standard form. The following data were recorded and included as many of the known prognostic factors as possible⁹: (1) age, sex, height and weight; (2) stated weight loss; (3) symptoms of cough, haemoptysis, shortness of breath, chest pain, other pain, malaise, anorexia, dysphagia, abdominal pain, hoarseness, headache and "other" (all scored using 0 = absent, 1 = mild, 2 = moderate, 3 = severe); (4) activity score (performance status) recorded using the WHO scale¹¹; (5) cell type (squamous, adenocarcinoma, large cell, bronchoalveolar cell, other or not known); (6) source of the specimen (brush, needle, pleural, fluid, and sputum for cytological examination and bronchial biopsy, needle biopsy, or pleural biopsy for histological examination); (7) haemoglobin, white blood cell count, lymphocyte count, erythrocyte sedimentation rate, alkaline phosphatase, aspartate aminotransferase, albumin, and plasma sodium levels; (8) extent of disease in the lung (recorded as endobronchial lesion, peripheral mass, hilar mass, lobar collapse/infiltration and pleural effusion on a four point scale); (9) spread within the chest (recorded as spread to the mediastinum, supraclavicular fossa, nodes, pericardium and chest wall and scored as 0 = absent, 1 = probable, 2 = definite); and (10) distant spread (recorded as spread to the brain, liver, bone, skin, and "other" and scored as 0 = absent, 1 = probable, and 2 = definite).

All the biochemical variables were taken from local laboratory results and a subset of patients had their results checked centrally in a reference laboratory.

Symptoms and their severity at the time of consultation were recorded by the examining doctor. The extent of the disease was assessed using the available clinical and radiological evidence and no attempt was made to impose a uniform assessment policy or programme of scanning.

At the time of the management decision the physician, usually a consultant but sometimes an experienced registrar or a senior registrar, indicated his or her opinion as to survival from that date in months. Patients were then managed as planned and attended for follow up visits at two monthly intervals until death or for a minimum of a year. No attempt was made to update physicians about the progress of the study during its completion. Permission for the study was given by the clinical ethics committees of the health authorities of the participating doctors. The dates of death for patients who were still alive at the end of the one year follow up period were obtained from the Yorkshire Cancer Registry database for the purpose of this analysis.

STATISTICAL ANALYSIS

Physicians' predictions versus actual survival

The accuracy of the physicians' predictions were compared with the actual survival of the patient using the following methods.

(1) The survival curves for actual survival and survival according to the physicians' predictions were produced.

(2) The difference between the physician's prediction and the actual survival of the patient was calculated for each patient as a measure of the accuracy of the prediction.

(3) Both the physician's prediction of survival and the actual survival of the patient were categorised as being ≤ 3 months, 3–9 months, and > 9 months. The sensitivity, specificity, and predictive values of the physicians' predictions were then calculated as was the efficiency.

Sensitivity was calculated as the percentage of patients in a specific survival category who were correctly predicted – that is, the number correctly predicted within that category/the total number of patients whose actual survival was within that category. The predictive value was calculated as the percentage of patients predicted by the physician as being within a specific survival category who were correctly predicted – that is, the number correctly predicted within that category/the total number of patients whose survival was predicted to be within that category. Specificity was calculated as the number of patients correctly classified as not belonging to a specific survival category/the number of patients wrongly classified as belonging to a specific survival category + the number of patients correctly classified as not belonging to a specific survival category. The efficiency of the physicians' opinions was calculated as the total percentage of patients correctly predicted.

Prognostic factor modelling

Correlation between variables was tested using Kendall's tau coefficient. If the coefficient was greater than 0.7 then the variables were said to be correlated.

All variables were first tested for prognostic significance in a univariate analysis using the log rank test.¹² Survival was measured from the time of presentation until death from any cause or the date of the last follow up visit. Variables found to be associated with survival duration in the univariate analysis, which were recorded for most of the patients, were included in a multivariate analysis of survival using the Cox's proportional hazards model.¹³ In a stepwise forward selection algorithm, each variable under consideration was sequentially tested for significant improvement in the predictive value of the model. An index was then obtained from the Cox's proportional hazards model including factors which were significantly associated with survival at the 5% level of significance. Patients were assigned into one of three prognostic groups on the basis of the presence of risk factors as identified by the Cox's model. Those patients with a prognostic index in the lower quartile were assigned to one group, those with a prognostic index in the upper quartile to another group, and those with a prognostic index within the interquartile range to a third group. The larger the value of the prognostic index, the worse was the prognosis for the patient. Thus, patients whose prognostic index was in the lower quartile are referred to as the good prognosis group, those in the upper quartile the poor prognosis group, and those in the interquartile range the moderate prognosis group.

The predictive power of the index was evaluated by plotting the Kaplan-Meier survival curves for each of the prognostic groups and noting the separation between the curves. Survival estimates and 95% confidence intervals were calculated at three, six, 12, and 24 months for each of the three groups.

All analyses were carried out using the statistical software package SAS version 6.10.

Results

Three hundred and thirty patients were studied. Of these patients, data from 122 were not used in the analysis (22 had benign disease, 11 small cell lung cancer, one pulmonary tuberculosis, two patients were entered twice (due to physician error), two patients had insufficient information provided, three were withdrawn by their physician for unknown reasons, and 81 were recommended at the initial visit to undergo curative treatment (64 surgery and 17 curative radiotherapy)). In addition, one patient who was diagnosed as having non-small cell lung cancer was excluded from the study population as he had data recorded only at the initial visit and was subsequently lost to follow up. Thus, all analyses were based on 207 patients for whom curative treatment was not recommended; 55 were to receive no immediate treatment, 105 to receive palliative radiotherapy, and 47 to receive some other therapy

Table 1 Baseline characteristics

Men	161 (78%)
Women	46 (22%)
Median age (range) in years	68 (40-84)
Spread	
Mediastinum	75 (37%)
Pericardium	5 (2%)
Chest wall	16 (8%)
Pleural effusion	34 (17%)
Not recorded	2
Activity score	
0, 1	115 (56%)
2, 3, 4	91 (44%)
Not recorded	1
Metastases	
No	143 (69%)
Yes	64 (31%)
Hoarseness	
Absent/mild	177 (86%)
Moderate/severe	30 (14%)
Malaise	
Absent/mild	96 (46%)
Moderate/severe	111 (54%)
Anorexia	
Absent/mild	94 (45%)
Moderate/severe	113 (55%)

Activity score 0 or 1 = normal or with effort.

Activity score 2, 3, 4 = restricted, confined or bedridden.

(chemotherapy, corticosteroids, new bronchodilators, antitussives, or analgesics).

Ninety eight patients (47%) had squamous cell carcinoma, 14 (7%) adenocarcinoma, 30 (14%) large cell carcinoma, one (< 1%) bronchoalveolar carcinoma, 10 (5%) had another type of carcinoma, and two (1%) did not have their cell type recorded. Fifty two patients (25%) were managed as non-small cell lung cancer without histological proof.

The baseline characteristics of these patients are presented in table 1. A detailed account of the patients' symptoms has been published elsewhere.¹⁴

Physicians' predictions compared with actual survival

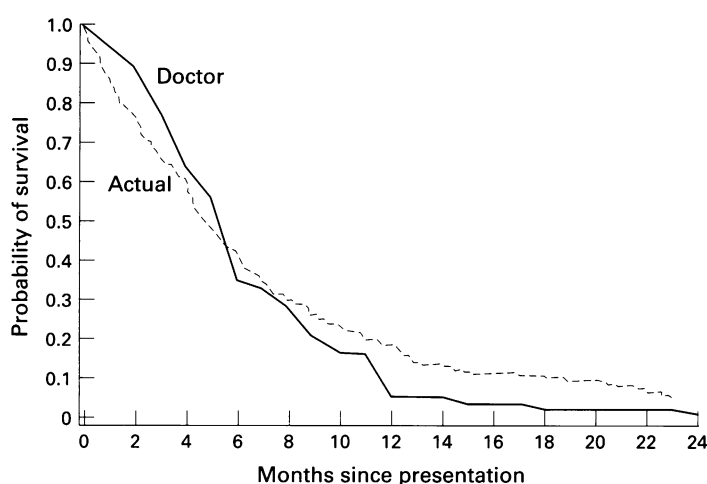
Predictions of how many months a patient would survive from the time of management decision and registration were made by physicians at that time for 203 of the 207 patients. Nine physicians in total examined patients in the study, but three recorded their opinions for just one patient so that, in practice, the predictions were derived from two clinical oncologists and four chest physicians.

Survival curves of physicians' predictions and actual survival

Figure 1 presents the survival curves for actual survival and survival according to the physicians' estimates. In general, physicians overestimated survival for patients who survived 0-5 months after their initial visit but, as actual survival increased, they were more likely to underestimate survival. This pattern of prediction for patients did not appear to vary between physicians.

Accuracy of physicians' predictions

The difference between the time of survival the physician predicted and the actual survival was calculated for the 196 patients whose date of death was known and who had their survival predicted. Three patients had their date of death recorded but did not have their survival



Number at risk												
Doctor	203	155	122	82	60	45	36	26	22	21	18	10
Actual	207	159	126	85	62	47	38	28	23	22	19	10

Figure 1 Actual survival and physicians' estimates.

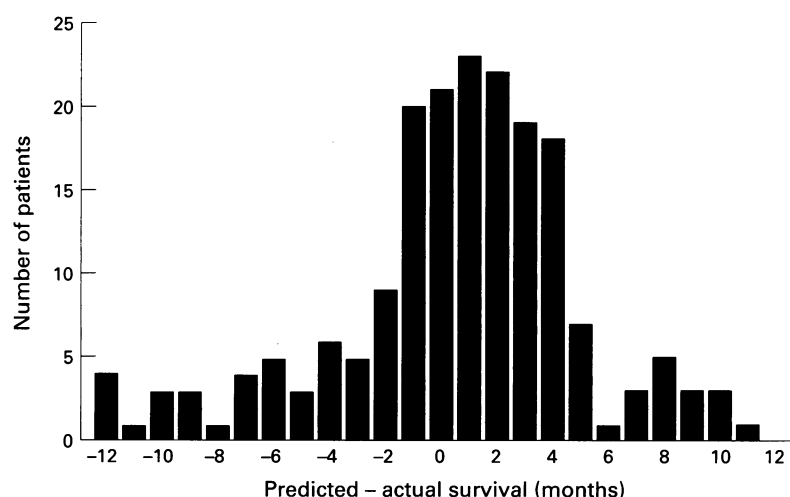


Figure 2 Accuracy of physicians' predictions.

predicted. Information about the date of death for eight patients was not available. Their survival times were censored at either their last follow up visit or the last time there was documented evidence that they were still alive. Thus, the difference between their predicted and actual survival could not be calculated (one of these patients did not have his survival predicted by his physician). One patient had his survival censored at nine months, four at 12 months, one at 13 months, one at 15 months, and one at 37 months. Thus, all we know about these patients is that they survived for at least nine, 12, 13, 15, and 37 months, respectively.

Table 2 Overall comparison of physicians' survival predictions and actual survival

Physician's prediction	Actual survival			Total
	≤ 3 months	3-9 months	> 9 months	
≤ 3 months	36	7	2	45
3-9 months	27	59	31	117
> 9 months	6	15	20	41
Total	69	81	53	203
	Sensitivity	Specificity	Predictive value	
≤ 3 months	52%	94%	80%	
3-9 months	73%	53%	50%	
> 9 months	38%	86%	49%	

Efficiency = 57%.

The accuracy of the physicians' predictions is presented in fig 2 which shows the differences between the physicians' predictions and the actual survival of patients. It includes data for the 184 patients whose predicted survival was known to be within 12 months of their actual survival. Nine patients had their survival underestimated by more than 12 months; these patients survived 16, 17, 19, 24, 25, 33, 36, 44, and 62 months longer than their physician had predicted. Three patients had their survival overestimated by more than 12 months; these patients died 14, 15, and 18 months earlier than predicted.

Of the 196 patients whose date of death was known and whose survival was predicted, 70 (36%) had their survival time underestimated and 107 (55%) had their survival time overestimated. Nineteen (10%) had their survival correctly predicted by their physician to within one month, 115 (59%) to within three months, and 139 (71%) to within four months of their actual survival.

Of the eight patients whose survival was estimated by the physician and who did not have a date of death recorded, four patients had their survival underestimated. Two patients who had their survival predicted as being 12 months were still alive at 12 months as was one patient whose survival had been estimated as 15 months. The remaining patient who had his survival censored at 15 months did not have a survival time predicted by his physician.

No obvious differences were noted between the performance of teams for which only a consultant gave a prognosis and others in which, on occasion, a senior registrar or a specialist registrar did so.

Categorisation of survival

The physicians in this study tended, in general, to give their patients with non-small cell lung cancer an estimate of their expected survival in terms of being three months or less, between three and nine months, and more than nine months. These categories were therefore adopted to summarise the predicted and actual survival of the patients.

Two hundred and three patients were included in one of the above categories. The eight patients who did not have a date of death recorded were all included in the category of having survived beyond nine months. The sensitivity, specificity, predictive values, and efficiency for the physicians' predictions into each of the three categories are presented in table 2.

Physicians predicted 45 patients to have a survival time of up to three months, of whom 36 (80%) died within three months. However, physicians failed to identify the remaining 33 patients (47% of the total) who did not survive beyond three months. Thus, there appeared to be a tendency to overestimate survival in these patients. So, although the predictive value was high, sensitivity was low. Few patients who were predicted to survive three months or less survived beyond three months. Thus, specificity was high (93%).

Of the 81 patients who actually survived for 3-9 months, physicians correctly predicted

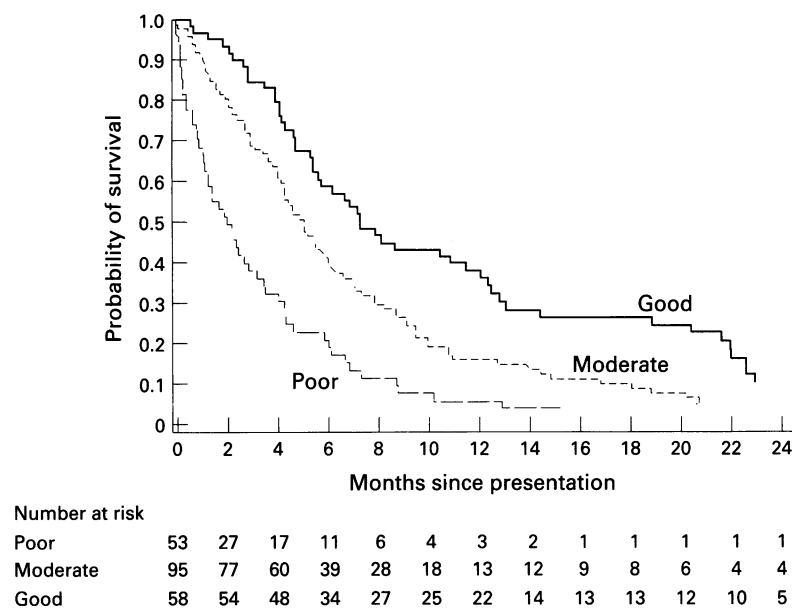


Figure 3 Model including sex, activity score, distant metastases, hoarseness, and malaise.

survival in 59 cases (73%). However, in addition to these 59 patients, they also predicted that 27 patients who actually survived three months or less, and 31 patients who survived beyond nine months, would survive for 3–9 months. The predictive value was therefore just 50%.

Of the 53 patients who survived more than nine months, only 20 patients (38%) were identified as such by physicians. In total, physicians had predicted that 41 patients would survive beyond nine months and thus were successful in predicting the survival of 49% of patients they classified. There appeared to be a tendency to underestimate the survival of patients who survived more than nine months.

Overall, 115 of 203 patients (57%) were correctly classified by physicians within the categories.

PROGNOSTIC FACTOR MODELLING

Using a prognostic index to define prognostic groups

Sex, activity score, the presence of distant metastases, hoarseness and malaise were identified

in the log rank test as being potentially important with respect to survival. These were included in the Cox model and the following prognostic index (PI) was derived:

$$PI = -0.49 \times \text{sex} + 0.35 \times \text{activity score} + 0.52 \times \text{distant metastases} + 0.57 \times \text{hoarseness} + 0.41 \times \text{malaise}.$$

Sex was coded as male (0) and female (1). The activity score was measured as being either normal (indicated on the data form as being normal or with effort: WHO score 0 or 1) or restricted (indicated on the data form as being restricted, confined or bedridden: WHO score 2, 3, or 4). These groups were coded as 0 and 1, respectively. Distant metastases were recorded as probable or definite (present, 1 or absent, 0). Symptoms of hoarseness and malaise were coded as being either absent (absent or mild, 0), or present (moderate or severe, 1).

Patients were assigned to the good prognosis group if their prognostic index was less than or equal to 0.41 (the lower quartile), and to the poor prognosis group if their prognostic index was greater than or equal to 0.93 (the upper quartile). If the prognostic index fell between these two values then patients were assigned to the moderate prognosis group.

Survival curves for each of the three prognostic groups are presented in fig 3. They show distinct separation between all three groups. Results of the log rank test and survival estimates and confidence intervals at three, six, 12, and 24 months are presented for each of the three groups in table 3.

It appeared, therefore, that the model identified three groups of patients with different survival rates. More patients in the poor prognosis group died at an earlier stage than in the good or moderate prognosis groups. It was more difficult to distinguish between the survival rates of patients in the good and moderate prognosis groups. Although the survival curves appeared to be separate, the confidence intervals at all time points examined overlapped. Using the model a physician could, however, determine whether a patient had a good, moderate, or poor prognosis and could estimate the prob-

Table 3 Summary of prognostic factor models

	A			B			C			D		
	Good	Moderate	Poor	Good	Moderate	Poor	Good	Moderate	Poor	Good	Moderate	Poor
No. of patients	58	95	53	51	100	51	57	73	73	46	74	41
χ^2 , df, p	35.94	2	< 0.001	58.78	2	< 0.001	40.65	2	< 0.001	49.95	2	< 0.001
Median survival (months)	7.3	5.0	2.1	8.2	4.9	1.4	8.2	5.7	1.7	8.4	4.5	1.9
Survival estimates (confidence intervals)												
3 months	84% (75 to 94)	71% (61 to 80)	38% (25 to 51)	90% (82 to 98)	70% (61 to 79)	31% (19 to 44)	84% (75 to 94)	81% (72 to 90)	36% (25 to 47)	89% (80 to 98)	73% (63 to 83)	32% (17 to 46)
6 months	59% (46 to 71)	41% (31 to 51)	21% (10 to 32)	65% (52 to 78)	41% (31 to 51)	16% (6 to 26)	63% (51 to 76)	45% (34 to 57)	18% (9 to 27)	67% (54 to 81)	38% (27 to 49)	20% (7 to 32)
12 months	38% (25 to 50)	16% (8 to 23)	6% (0 to 12)	43% (30 to 57)	14% (7 to 21)	4% (0 to 9)	39% (26 to 51)	19% (10 to 28)	5% (0 to 11)	39% (24 to 53)	20% (11 to 29)	0% (0 to 0)
24 months	10% (2 to 18)	5% (0 to 5)	0% (0 to 0)	5% (1 to 20)	5% (1 to 9)	0% (0 to 0)	8% (1 to 16)	9% (2 to 16)	3% (0 to 6)	14% (3 to 25)	6% (1 to 12)	0% (0 to 0)

A = model including sex, activity score, distant metastases, hoarseness and malaise.

B = model including sex, activity score, distant metastases, hoarseness and malaise and physicians' prediction of survival.

C = model including only the physicians' predictions.

D = model including distant metastases, hoarseness, malaise, treatment intent, lymphocyte count, albumin, sodium, and alkaline phosphatase concentration.

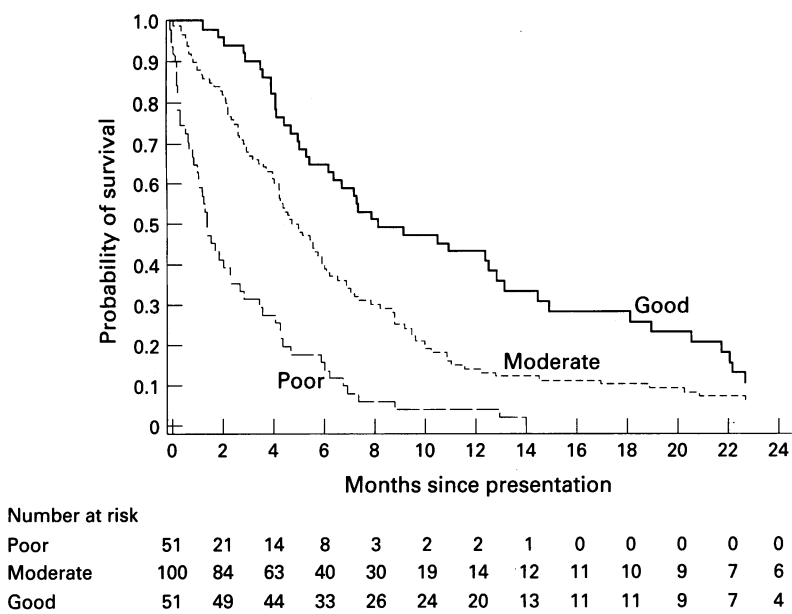


Figure 4 Model including sex, activity score, distant metastases, hoarseness, malaise, and physicians' predictions of survival; $p = 0.0001$ (log rank test).

ability that a patient would still be alive at a particular time point between the initial visit and 24 months.

Model including the physician's prediction of survival

In order to investigate how much more (if any) information about the prognosis of the patients the physicians' predictions gave, the actual prediction made by the physicians in months was forced into the Cox's proportional hazards model. The prognostic index was then calculated as follows:

$$\text{PI} = -0.51 \times \text{sex} + 0.17 \times \text{activity score} + 0.35 \times \text{distant metastases} + 0.60 \times \text{hoarse} + 0.41 \times \text{malaise} - 0.07 \times \text{physician's prediction of survival (months)}.$$

Patients were assigned to the good prognostic group if their prognostic index was less than or equal to -0.56 (the lower quartile), to the poor prognostic group if their index was

greater than or equal to 0.36 (the upper quartile), and to the moderate prognosis group if the prognostic index fell between these values. Survival was then plotted for the three groups identified by this partitioning (fig 4) and survival estimates and 95% confidence intervals at three, six, 12, and 24 months calculated as above (table 3).

Again, the survival curves for the three groups were well separated. The difference between the survival curves for the poor prognosis group and the moderate and good prognosis groups appeared to be more marked in the early months than with the first model. This would be expected as the physicians' predictions were shown to be highly specific for patients who survived three months or less. It appeared that the poor prognosis group identified using this model contained patients with a shorter survival than those in the group identified using the model which did not include the physicians' predictions. It was difficult to distinguish between the moderate and good prognosis groups in the early months but at 12 months there did appear to be a difference in the survival rates in each of the three groups. By 24 months no patients in the poor prognosis group were alive.

Model including only the physicians' predictions

In order to compare the outcomes achieved using the prognostic factor model with those achieved using the predictions made by the physicians, a model using the time the physician predicted a patient would survive as the sole prognostic factor in the Cox's proportional hazards model was developed to give the following prognostic index:

$$\text{PI} = -0.096 \times \text{physician's prediction of survival (months)}$$

Patients were assigned to the good prognosis group if their prognostic index was less than or equal to -0.87 , to the poor prognosis group if their prognostic index was greater than or equal to -0.38 , and to the moderate prognosis group if their prognostic index was between these values. Again, the survival curves were plotted for these three prognostic groups (fig 5) and survival estimates and 95% confidence intervals calculated at three, six, 12, and 24 months as above (table 3).

The survival curves for the groups again showed good separation. As in the previous models, at three months there was a distinct separation between the poor prognosis group and the good and moderate prognosis groups. However, the separation between the moderate and good prognosis groups was not so distinct as that observed using the prognostic factor model. It appeared that the model which included the prognostic factors was able to distinguish between the good and moderate prognosis patients earlier than the model including only the physicians' opinions.

Model including lymphocyte count and serum albumin concentration

In previous trials in small cell lung cancer some laboratory values have been shown to be significant prognostic factors¹⁵ and it was

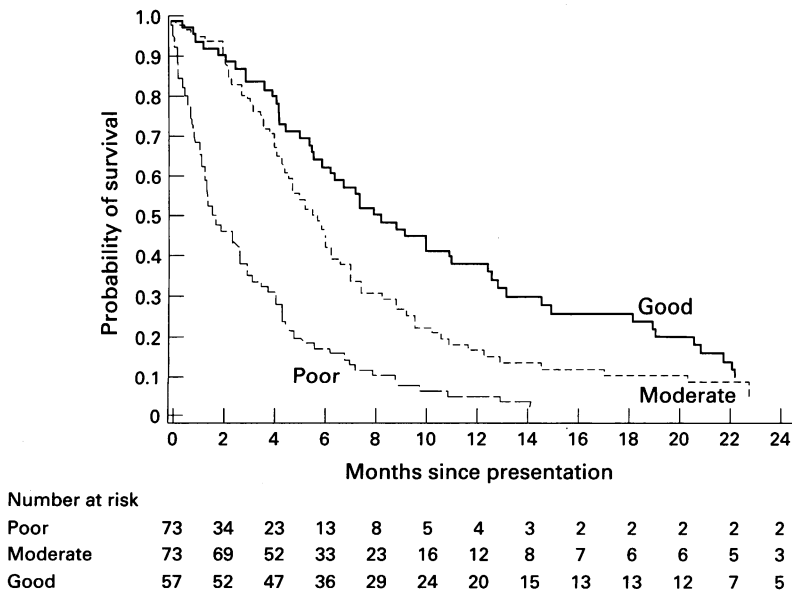


Figure 5 Model including only physicians' predictions of survival; $p = 0.0001$ (log rank test).

decided to examine laboratory values for significant correlations with survival at 12 months. The log rank test indicated that lymphocyte count, and albumin, sodium and alkaline phosphatase concentrations were possibly important prognostic factors. These were recorded at presentation for 160 patients (77%) and were included in the Cox model to determine if they were prognostic factors of significance.

The prognostic index was calculated as follows:

$$\text{PI} = -0.42 \times \text{distant metastases} + 1.1 \times \text{hoarseness} + 0.47 \times \text{malaise} - 0.34 \times \text{immediate treatment intent} + 0.72 \times \text{lymphocyte count} + 0.94 \times \text{serum albumin} + 0.62 \times \text{sodium} - 0.98 \times \text{alkaline phosphatase}$$

Immediate treatment intent was coded as 1 for supportive care, 2 for non-curative radiotherapy, and 3 for some other undefined non-curative treatment. Lymphocyte count was coded as 1 for a count $< 1 \times 10^9/\text{l}$ and 0 for a count $> 1 \times 10^9/\text{l}$. Concentrations of serum albumin, sodium, and alkaline phosphatase were coded as 1 for $< 30 \text{ g/l}$ and 0 for $> 30 \text{ g/l}$, 1 for $< 135 \text{ IU/l}$ and 0 for $\geq 135 \text{ IU/l}$, and 1 for normal and 0 for abnormal, respectively.

The patients were assigned to the good prognosis group if their prognostic index was less than or equal to -2.85 , to the poor prognosis group if their prognostic index was greater than or equal to -1.6 , and to the moderate prognosis group if their prognostic index was between these values.

Survival estimates and 95% confidence intervals were calculated at three, six, 12, and 24 months (table 3). The survival curves for the three groups appeared to be clearly separated (fig 6). The poor prognosis group did not contain any patients who were alive at nine months. The median survival for the good prognosis group was 8.4 months, implying that the inclusion of these laboratory measures produced a model which had a good prognosis

group with a longer expected survival than the good prognosis groups produced by the original model where the median survival was 7.3 months. However, the number of patients contributing to this model was 23% smaller than that of the original prognostic model.

Incorporation of the physicians' predictions of survival in months to this model containing the laboratory variables appeared to improve discrimination between the poor and the moderate and good prognosis groups further.

COMPARISON OF MODELS

Physician's opinion versus prognostic factor models
No formal comparisons were made between the models. However, it was clear from the survival curves that the model which included only the physicians' predictions of survival was less successful in distinguishing between the moderate and good prognosis groups in the initial months than the prognostic factor models.

However, there was little difference in the survival curves of the poor prognosis groups defined by any of the models. The best discrimination between the poor and moderate/good prognosis groups was provided by the model which included prognostic factors incorporating the physicians' predictions of survival. Thus, it would appear that physicians are not necessarily using the same factors as the model to predict survival and must be using additional information on which to base their estimates of survival. The prognostic factor model which included the laboratory data at presentation provided the best separation between groups. It must be emphasised, however, that fewer patients ($n = 160$) contributed to this model.

Discussion

This study population was unselected, sequential, and comparable to two large regional UK surveys in its distribution of age, sex, and histological type.^{2,4}

The prognoses of patients with lung cancer represent a continuum, ranging from those patients who can almost certainly be cured, to those who are certainly incurable. In this study only patients not being considered for curative treatment were included.

The participating physicians were not aware, in any formal sense, of the weightings of the different variables they were recording. They gave optimistic prognoses to many patients whose survival was limited to less than three months (table 2), and this optimism seemed to extend to five months (fig 2). Conversely, however, a physician giving a prognosis of less than three months was likely to be correct. In this study 78% of patients identified as having survival of three months or less did, indeed, die within three months. Evans and McCarthy¹⁶ found that similar optimistic predictions were made by a terminal care support team who made multiple visits to 42 cancer patients at home. Better correlations with survival were observed by estimates of the Karnofsky score at the visits. Doctors might tend to give more optimistic prognoses to many patients for very

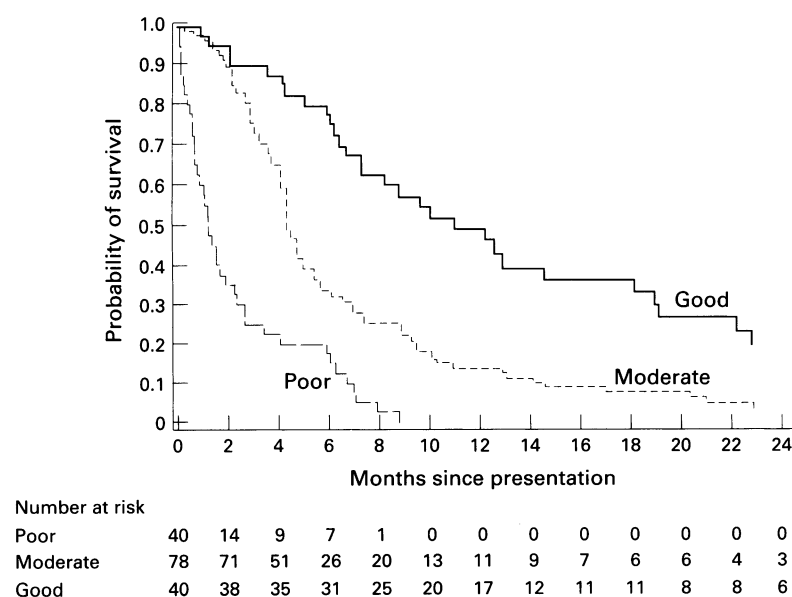


Figure 6 Model including distant metastases, hoarseness, malaise, treatment intent, lymphocyte count, albumin, sodium and alkaline phosphatase concentration, and physicians' predictions of survival; $p = 0.0001$ (log rank test).

understandable psychological reasons. This study, however, suggests that a validated prognostic factor model might be able to assist in the recognition of other patients with a short life expectancy.

It must be emphasised that categorising the physicians' predictions into < 3 months, 3–9 months, and > 9 months, and comparing this with actual survival, may not be a true representation of the physicians' accuracy since information is lost by categorising. Also, the physicians were not asked to classify the survival of the patient into one of these three categories. If asked to do so they may have recorded a different prognosis.

The major prognostic factors in lung cancer have been recognised for several years^{9 17–19} and numerous clinical items are known to be significantly correlated with survival. For example, Stanley⁹ found that, of the 78 factors analysable from data on 2671 patients entered into a Veterans Administration Lung Group trial, 23 had a *p* value of < 0.0001. In previous studies^{17 20–22} performance status (activity score) had a very strong individual correlation with survival (Cox analysis), while disease extent and a measure of the anorexia/malaise/weight loss complex were other powerful variables. The laboratory values of strongest significance were a low lymphocyte count¹⁷ and albumin concentration.²³ When these variables were included in our multivariate analysis they were shown to contribute to a model from which a prognostic index could be derived.

This study confirmed the findings in previous studies that activity score (performance status), disease extent, and malaise symptoms were associated with survival. It also showed that, for this study population, the sex of the patient and the symptom of hoarseness were significant prognostic indicators. Although 23% of patients did not have complete laboratory data, it was evident that lymphocyte count and albumin, sodium, and alkaline phosphatase concentrations at presentation were significant prognostic indicators.

As well as allowing stratification of patients for clinical trials, prognostic factor models should, in theory, be able to improve a physician's ability to manage his or her patients. This would undoubtedly be useful, not only to avoid giving incorrect prognoses to patients and their relatives, but also to allow a more rational choice of treatment. In practice it is unlikely that prognostic factor models would on their own be clinically useful unless they were very accurate.

A surrogate study with 26 consultant members of our collaborative group, the Yorkshire Thoracic Group, showed that, for a patient estimated by the doctor to survive more than three months, all of the doctors would change their management policy if a prognostic factor model predicted survival at less than three months with a 90% accuracy, but only 12 (46%) would do so if the probability of the accuracy of the model was only 70%. The same exercise for six month survival gave similar figures of 16 (62%) and eight (31%), respectively.

Accordingly it seems that, for a model to influence the management decisions of physicians based on their own appreciation of the clinical picture, it would have to have a predictive power of 80% or better. However, in this study the physician's overall accuracy in (indirectly) classifying patients into one of three categories was far below this figure at 57%.

All the models yielded prognostic groups which were distinct in terms of survival. All were similar in that they identified a poor prognosis group which was distinct from the moderate and good prognosis groups from the start of the study. The model which included laboratory variables was the most successful in separating a moderate and a good prognosis group. Indeed, to achieve better discrimination between patients with a poor, moderate, and good prognosis it appeared from this study that both physicians' opinions and prognostic factor information were necessary. In all other models, although there did appear to be good separation between the curves, at all time points until 12 months the confidence intervals for survival overlapped.

From the information collected in this study only patients with a poor prognosis could be identified with any degree of precision, both using the physicians' predictions and the prognostic factor model. Perhaps further information collected later in a patient's disease progression would allow a better estimate of how long a patient who did not have a poor prognosis would survive. Perhaps, too, some other baseline laboratory measures (not included in our prognostic factor model) would have been significantly associated with survival and would have allowed further discrimination between the three prognostic groups.

The prognostic models as developed in this analysis clearly require refinement and further study before they could be used to improve a physician's prediction of a given patient's prognosis. Thus, a further study using more baseline laboratory variables would be worthwhile and, in addition, the importance of lymphocyte count, serum albumin, sodium, and alkaline phosphatase concentration could be investigated more fully. Such a study could also investigate the importance as potential prognostic factors of other variables, as yet unidentified, which a physician might take into account when making an estimate of how long a patient would survive. If a further study was planned, it should aim to collect more information from more patients. Not only could this dataset then be used to refine models developed in this study but, if large enough, it could be used to develop and cross validate any new models which might come to light following the measurement of additional variables.

The authors wish to thank the Yorkshire Clinical Trials and Research Unit for their invaluable support and Mrs E Power for her secretarial assistance. This study was generously supported by the Yorkshire Cancer Organisation.

The following consultants and their colleagues participated in the study: D Ash, radiotherapist (Leeds/Huddersfield); H J Close, radiotherapist (Leeds/Halifax); C K Connolly, physician (Northallerton and Darlington); M F Muers, physician (Leeds); M D Peake, physician (Pontefract); S B Pearson, physician (Leeds); Professor E H Cooper (University of Leeds); Mrs M Jones, statistician (YCO, Leeds); Miss P Shevlin, statistician (YCO, Leeds); Mrs J Brown, statistician (YCO, Leeds).

- 1 Lee PN, Fry JS, Forey B. A. Trends in lung cancer, chronic obstructive lung disease and emphysema death rates for England and Wales 1941-1985 and their relation to trends in cigarette smoking. *Thorax* 1990;45:657-65.
- 2 Connolly CK, Jones WG, Throgood J, Head C, Muers MF. Investigation, treatment and prognosis of bronchial carcinoma in the Yorkshire Region of England, 1976-1983. *Br J Cancer* 1990;61:579-83.
- 3 Watkin SW, Hayhurst GK, Green JA. Time trends in the outcome of lung cancer management: a study of 9090 cases diagnosed in the Mersey Region 1974-1986. *Br J Cancer* 1990;61:590-6.
- 4 Capewell S for the Edinburgh Lung Cancer Group. Patients presenting with lung cancer in South East Scotland. *Thorax* 1987;42:853-7.
- 5 Spiro SG. The staging of lung cancer. In: Flenley DC, Petty TL, eds. *Recent advances in respiratory medicine* 4. Edinburgh:Churchill and Livingstone,1986:261-75.
- 6 Mountain CF. A new international staging system for lung cancer. *Chest* 1986; 89(Suppl):225S.
- 7 Spiro SG. Lung tumours. In: Brewis RAL, Gibson JG, Geddes DM, eds. *Respiratory medicine*. London:Balliere Tindall,1990:832-79.
- 8 Rawson NSB, Peto J. An overvoew of prognostic factors in small cell lung cancer. *Br J Cancer* 1990;61:597-604.
- 9 Stanley KE. Prognostic factors for survival in patients with inoperable lung cancer. *J Natl Cancer Inst* 1980;65:25-32.
- 10 Cooper EH, Splinter TAW, Brown DA, Muers MF, Peake MD, Pearson SB. Evaluation of a radioimmunoassay for neuron specific enolase in small cell lung cancer. *Br J Cancer* 1985;52:333-8.
- 11 World Health Organisation. *Handbook for reporting results of cancer treatment*. Offset publication 48.10. Geneva: WHO,1979.
- 12 Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, *et al*. Design and analysis of randomised clinical trials requiring prolonged observation of each patient. *Br J Cancer* 1976;34:585-612;35:1-38.
- 13 Cox DR. Regression models and life tables (with discussion). *J R Stat Soc B* 1972;34: 187-90.
- 14 Muers MF, Round CE. Palliation of symptoms in non-small cell lung cancer: a study in the Yorkshire Regional Cancer Organisation thoracic group. *Thorax* 1993;48:339-43.
- 15 Souhami RL, Bradbury I, Geddes DM, Spiro SG, Harper PG, Tobias JS. Prognostic significance of laboratory parameters measured at diagnosis in small cell cancer of the lung. *Cancer Res* 1985;45:2878-82.
- 16 Evans C, McCarthy M. Prognostic uncertainty in terminal care: can the Karnofsky index help? *Lancet* 1985;i:1204-6.
- 17 Vincent MD, Ashley SE, Smith LE. Prognostic factors in small cell lung cancer: a simple prognostic index is better than conventional staging. *Eur J Cancer Clin Oncol* 1987;23:1589-99.
- 18 Stanley KE. Prognostic factors in lung cancer. In: Aisner J, ed. *Lung cancer*. New York: Churchill Livingstone,1985:41-66.
- 19 Lagakos S. Prognostic factors for survival time in inoperable lung cancer. In: Strauss MJ, ed. *Lung cancer: clinical diagnosis and treatment*. New York:Grune and Stratton, 1977:271-80.
- 20 O'Connell JP, Kris MG, Gralla RJ, Groshen S, Trust A, Fiore JJ, *et al*. Frequency and prognostic importance of pre-treatment clinical characteristics in patients with advanced non-small cell lung cancer treated with combination chemotherapy. *J Clin Oncol* 1986;4:1604-14.
- 21 Hyde L, Wolf J, McCracken S, Yesner R. Natural course of inoperable lung cancer. *Chest* 1973;64:309-12.
- 22 Capewell S, Sudlow M, on behalf of the Edinburgh Lung Cancer Group. Performance and prognosis in patients with lung cancer. *Thorax* 1990;45:951-6.
- 23 Marshall RJ, Curzon PDG, Pearson SB, Cooper EH, Muers MF, Peake MD. Prognosis in squamous cell lung cancer. The contribution of plasma proteins. *Tumour Diagn Ther* 1985;5:195-8.