Pergamon

European Journal of Cancer Vol. 32A, No. 7, pp. 1135–1141, 1996

Printed in Great Britain
0959-8049/96 $15.00+0.00

0959-8049(95)00664-8

Original Paper

Karnofsky and ECOG Performance Status Scoring in Lung Cancer: A Prospective, Longitudinal Study of 536 Patients From a Single Institution

G. Buccheri, D. Ferrigno and M. Tamburini

Lung Cancer Unit, 'A. Carle' Hospital, Cuneo, and the Division of Psychological Researches, National Tumour Institute, Milan, Italy

The Karnofsky's index of performance status (KPS) and the Eastern Cooperative Oncology Group Performance Status Scale (ECOG PS) are widely used methods of assessing the functional status of cancer patients. In this study, we compare their predictive validity, and suggest a table of transformation between scales. 536 consecutive lung cancer patients were assigned both KPS and ECOG PS scores before, during and after treatment (in all, 1656 assignments). Patients were accurately staged at diagnosis, and carefully re-evaluated at each follow-up visit. Multiple clinical, laboratory and instrumental data were recorded along with performance status assessments. Survival times were measured from the pathological diagnosis. KPS and ECOG PS assignments were strongly related to each other (Spearman $R = -0.869$). Correlation between scales persisted unchanged in pretreatment and post-treatment assessments, advanced and limited diseases, response or non-response to treatment, and different assessors (R indices ranging from –0.825 to –0.901). A three-point conversion table showed the highest rate of success with an overall percentage of agreement exceeding 84% (grade 1: KPS = 100, 90, 80 and ECOG PS = 0, 1; grade 2: KPS = 70, 60 and ECOG PS = 2; grade 3: KPS < 60 and ECOG PS = 3, 4). Both univariate and multivariate analyses of survival documented the predictive validity of the two scales. However, KPS showed less ability than ECOG PS to discriminate patients with different prognosis. Because of the better predictive ability shown in this study, ECOG PS should be preferred to KPS. A general consensus on the scale to use could avoid problems of conversion, which is not always easy and free of errors.

Key words: bronchogenic carcinoma, performance status scoring, scale equivalence, predictive validity
Eur J Cancer, Vol. 32A, No. 7, pp. 1135–1141, 1996

INTRODUCTION

PERFORMANCE STATUS (PS) is a global assessment of the patients' actual level of function and ability of self-care. PS is a major prognostic factor, a predictor of the benefit and toxicity of treatments, as well as an indicator of comorbidity and other 'host factors' [1]. PS has been applied in cancer research to select and stratify patients for inclusion in treatment trials [2], to measure the efficacy of treatment [3], to assess the 'quality of survival' of cancer patients [4] and to estimate prognosis [5]. In a recent review, focusing on this last point [6], PS was prognostically important, or very important, in each of the studies mentioned (only three investigations, performed on particularly selected patients, did not confirm this factor as a significant covariate for multivariate models of survival).

Several measures of PS are available for use [2], among them the Karnofsky's Scale of Performance Status (KPS) and Eastern Cooperative Oncology Group Scale of Performance Status (ECOG PS). The Karnofsky index was introduced in the 1940s, at the beginning of cancer chemotherapy [7]. It is rated on a scale of 0–100, in steps of 10. It describes the patient's ability to perform normal activity and do active work, and whether there is any need for assistance. At 100 all is well, at 0 the patient is dead. KPS has proved useful and has

Correspondence to G. Buccheri, via Republica 10/c, 1-12018 Roccavione (CN), Italy.
This paper was presented in part at the 1994 ASCO Annual Meeting, 14–17 May, Dallas, Texas, U.S.A.
Received 21 Aug. 1995; revised 20 Nov. 1995; accepted 24 Nov. 1995.

survived in both practical and scientific oncology. In 1960, an alternate categorisation of PS was developed by the Eastern Cooperative Oncology Group [8]. The ECOG PS provides a five-point scale in contrast to the 11-point Karnofsky scale. It is simpler to use and has a precise but narrow message.

At present, both scales are very popular and widely used. In a recent Medline search, aiming to ascertain which scale of PS was the most frequently used [9], 114 authors were found to make use of KPS, while 113 employed ECOG PS. There are no obvious reasons to prefer one scale or another. Both scales are moderately valid and reliable [10–13], while there are no comparative studies.

This study was mainly designed to compare the prognostic significance of KPS and ECOG PS. A secondary aim was to obtain a conversion table between scales, looking at the possible factors which may affect the equivalence. To this end, we report a large, homogeneous cohort of longitudinally studied patients (lung cancer patients), seen consecutively, during the last 6 years, in a single institution.

PATIENTS AND METHODS

Patients

Between September 1988 and August 1994, 536 new patients with a cytologically or histologically proven bronchogenic carcinoma were seen at the A. Carle Hospital of Chest Diseases. Males constituted 90% (480/536) of the cohort studied. The median age was 65 years (range 32–87 years). All patients, except 40, were currently smokers or ex-smokers (median pack-year = 46, range 0–230). Median diagnostic delay (from the first warning symptom to the histopathological diagnosis) was 3.1 months (range 0–36 months). Median weight loss in the 6 months preceding the diagnosis was 3% (range 0–38%). All histopathological diagnoses were carried out, in accordance with the revised World Health Organisation (WHO) classification of lung tumours [14], by one pathologist (S.R.). The distribution of tumour cell types was as follows: squamous cell carcinomas (243 patients, 45%), adenocarcinomas (88, 16%), small cell cancers (67, 13%), large cell anaplastic carcinomas (33, 6%), anaplastic undefined or mixed adenosquamous carcinomas (105, 20%). Simple biochemistry tests, haematological counts and tumour markers (carcinoembryonic antigen (CEA), and tissue polypeptide antigen (TPA) [15, 16]) were obtained in all patients at diagnosis and then at each follow-up evaluation. Pretreatment diagnostic and staging procedures routinely included physical examination, chest radiographs, bronchoscopy, bronchial biopsies and/or cytological samplings. Computed tomography of the thorax, brain and superior abdomen has become a routine evaluation procedure for all patients, during the last 2–3 years, but, at the beginning of the study, it was performed only in patients considered for surgery. Almost all surgical candidates were examined by an additional total body immunoscintigraphy with anti-CEA monoclonal antibodies, according to a different, specific protocol [17, 18]. Ultrasound and other radioisotope scans were performed as clinically indicated. 114 patients underwent thoracotomy or other forms of pathological staging, such as mediastinoscopy, mediastinotomy and biopsies of distant metastatic deposits. On the basis of the information gained, patients were classified using the new International Union Against Cancer (UICC) staging system [19]. Staging classification included: 2 stage 0 (*in situ* carcinomas, pathological diagnosis), 59 stage I (11%), 52 stage II (10%), 96 stage IIIa (18%), 146 stage IIIb (27%) and 181 stage IV patients (34%).

94 individuals (18%) underwent surgical resection, 57 (11%) had radiotherapy as primary treatment and 205 (38%) were treated by chemotherapy (the MACC regimen, with or without lonidamine, in patients with non-small cell lung cancer [20, 21]; and a variety of different individualised chemotherapy programmes in the smaller group of small-cell lung cancer patients). The remaining 180 patients (34%) had only palliative irradiation or supportive medical care. No standard treatment was planned in case of recurrence or progression of disease; patients were given the most suitable therapy for their state.

Follow-up consisted of clinical, laboratory and radiographic re-assessments. They were repeated monthly (or every 3 weeks) during chemotherapy, radiotherapy and medical symptomatic care. After a successful surgical tumour resection, follow-up visits were scheduled every 3 months during the first 2 years, and then twice a year. Patients were re-classified according to their actual stage of disease and treatment response category (as defined by Miller and associates [22]). For a variety of reasons, including a rapidly progressing disease, 282 patients were unable to start a minimum follow-up programme. Others abandoned their follow-up project some time after the diagnosis. In all, the number of follow-up visits per patient ranged from 0 to 18. The total number of clinical re-evaluations performed at each follow-up time (time 1 up to time 18) was 254, 195, 159, 132, 99, 76, 58, 42, 30, 25, 19, 10, 8, 4, 4, 3, 1 and 1, respectively.

All patients were assessable for survival analysis. This was accomplished through a telephone interview with the patient him- or herself, the patient's relatives or the referring community doctor. In this way, no patient was lost to follow-up and all survival times were updated to October 1994.

Table 1 summarises the main clinical characteristics of the cohort studied.

Performance status assessments

PS score was rated at diagnosis and during each follow-up visit. KPS and ECOG PS scores were assessed at once, and then recorded, along with other clinical, laboratory, instrumental and pathological data concerning the same patient, in a computer database file. Only doctors (the authors) were in charge of PS ratings. A brief interview specifically addressed to select the more appropriate categories of performance status preceded, as a routine clinical approach to the patient, the inquiry for symptoms and the physical examination.

Statistical analysis

Survival duration was measured in all patients from the time of pathological diagnosis. Univariate survival tests were performed using multiple grouping choices, with the specific purpose of securing an equal number of groups, and possibly similar numbers of patients per group, in both KPS and ECOG PS evaluations. Only performance status assessments at the initial clinic visit were used in this type of analysis. Survival curves were obtained using the Kaplan–Meier product-limit estimates of the survival distribution [23]. Differences between survival probabilities were tested by the log-rank test [24]. Multivariate survival analyses were accomplished using the Cox' proportional hazards regression model [25]. Two kinds of models were tested: (i) models containing both KPS and ECOG PS in addition to other

Table 1. Anthropometric, clinical, haematological, biochemical and pathological characteristics of the cohort studied

Variable	No.	Median/frequency	Lower quartile	Upper quartile
Age (years)	536	65	58	70
Sex (male/female)	536	480/56		
PBW (%)	530	97	92	100
Karnofsky PS (100–40)	536	31/96/139/171/64/29/6		
ECOG PS (0–4)	536	55/211/214/51/5		
Erythrocyte sedimentation rate	513	35	18	64
Haemoglobin (g/dl)	533	14	13	15
White cell count (no./mm^3)	533	8661	7100	11190
Neutrophils (no./mm^3)	530	5880	4580	7985
Platelets (no./mm^3)	530	294000	233000	372000
Total serum proteins (g/dl)	527	7	6	7
Lactate dehydrogenase (g/dl)	520	325	271	420
Alkaline phosphatase (g/dl)	530	181	128	237
sGOT (g/dl)	532	18	14	25
sGPT (g/dl)	532	16	11	26
Creatininie (g/dl)	532	1	1	1
Uric acid (g/dl)	511	5	4	6
Sodium (mEq/l)	525	140	138	142
Carcinoembryonic antigen (ng/ml)	496	3	1	7
Tissue polypeptide antigen (U/l)	491	150	80	280
T factor (0–4)	536	1/64/194/87/190		
N factor (0–3)	536	163/55/199/119		
M factor (0–1)	536	355/181		
Clinical stage (0–4)	536	2/59/52/96/146/181		
Cell type (E/A/S/L/ND/M)	536	243/88/67/33/97/8		
Treatment instituted (P/C/R/S)	536	180/205/57/94		

PBW, percentage of the body weight 6 months before the diagnosis; PS, performance status; ECOG, Eastern Cooperative Oncology Group; sGOT, serum glutamic oxalacetic transaminase; sGPT, serum glutamic pyruvic transaminase; E, epidermoid; A, adenocarcinoma; S, small cell; L, large cell; ND, undefined; M, mixed adeno-squamous; P, palliative care, C; chemotherapy; R, radiotherapy; S, surgery.

prognostically meaningful variables (all the available ones with less than 15% of missing values and, on the other end of the spectrum, only variables used customarily for clinical trial stratification); and (ii) models containing either KPS or ECOG PS. For models with several covariates, a stepwise backward selection procedure was followed, in order to specify a model where all parameters had statistical significance ($P < 0.05$).

Given the ordinal character of both KPS and ECOG PS, the non-normality of their distribution (Shapiro–Wilk $W = 0.922$ and 0.227, $P = 0.000$), the non-equality of variance of KPS ($P < 0.001$, Levene test), and the failure to recuperate a normality pattern using common data transformations, we used only Spearman rank tests to estimate the correlation between PS scores [26].

All cited P values correspond to two-tailed significance tests. Data were processed using the STATISTICA package (StatSoft, Inc. 1993) running on a personal computer.

RESULTS

536 consecutive patients were scored by both KPS and ECOG PS scales during the 6 years of study. Nearly half could be re-assessed at least once during their treatment and/or post-treatment observation period. In all, 1656 assessments (1096 performed by one doctor, G.B., 560 by the other, D.F.) were available for evaluation. The distribution of KPS and ECOG PS total counts is displayed in Figure 1. At the closure of the study, 142 patients were still alive, 394 (74%) were already dead. The median survival (Kaplan–Meier estimate) for the

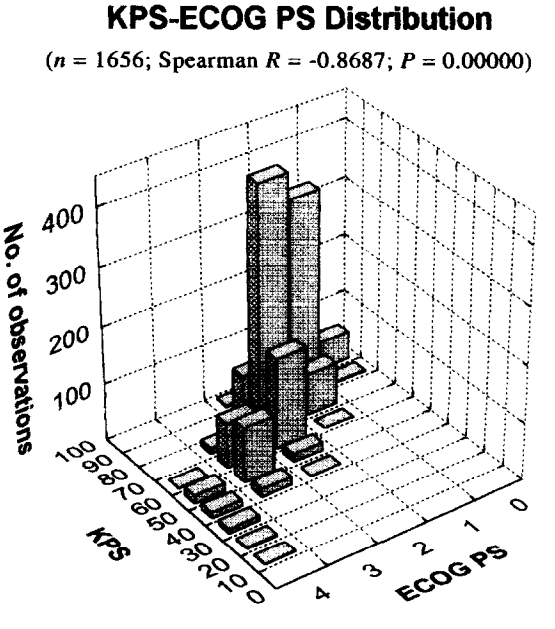

Figure 1. Distribution of Karnofsky's index and Eastern Cooperative Oncology Group performance status scoring (in all, 1656 paired assignments).

whole group was 8 months (cumulative proportion of survival at 8 months = 0.502, S.E. = 0.0227).

Survival analyses

Performance status, whatever the scale and the coding criteria used, was, as expected, highly significant in each of the analyses performed.

Using the equal size criterion, two categories of performance status came out as the best for a homogeneous comparison between scales. For a fortunate accident, the number of patients with a pretreatment ECOG PS = 0, 1 matched the number of patients with pretreatment KPS = 100, 90, 80 (266 subjects in all). The same occurred for patients with low PS (270 cases had ECOG PS = 2–4 or KPS <80). This circumstance permitted a comparison among scales in which only their capability of differentiating patients with different prognosis was the source of findings. As shown in Figure 2, the separation of the survival curves and the log-rank test statistic were considerable for both ECOG PS and KPS, but were more important for ECOG PS (–7.073 versus –6.086, log-rank statistic).

A backward stepwise selection was started from a proportional hazards regression model of 25 variables recorded at diagnosis (i.e. sex, age, percentage of the body weight 6 months before diagnosis (PBW), KPS, ECOG PS, tumour cell type (squamous versus non-squamous), clinical stage T, N and M factors, plus the 15 haematological, biochemical and radioimmunological parameters listed in Table 1). The stepwise procedure ended with the final model shown in Table 2, where ECOG PS was the most significant covariate, preceding the parameters of disease extension (T, N, M factors, and stage), PBW, white cell counts (neutrophils and total white cells), sex, LDH and TPA serum concentration. KPS was excluded by the model. A second run of Cox's multivariate analyses was launched using only those variables, which are normally incorporated into stratification trees (i.e. PBW, performance status—either both of the two scales or only one—and clinical stage). In this type of analysis, ECOG PS also exhibited a clear superiority in terms of prognostic capability and predictive validity (Table 3). In fact, when both scales were included among the explanatory co-variates, the statistical significance of ECOG PS surpassed that of KPS (t value: 1.70 versus –1.2). More importantly, when only one PS scale was incorporated, the model containing the ECOG PS scale was clearly more predictive (overall $\chi^2 = 156.75$ versus 116.25).

Correlation between scales and conversion tables

As shown in Figure 1, the overall degree of correlation between ECOG PS and KPS scoring was high (Spearman $R = -0.869$, 1656 observations). In particular, the Spearman rank correlation indices were: –0.858 at diagnosis, and –0.894, –0.866, –0.825, –0.865 and –0.901 at the first, second, third, fourth and fifth follow-up evaluations (numerically the most common follow-up reassessments). In addition, in the subgroups of patients with opposite stages of disease (stage I–IIIa versus stage IIIb and IV), as well as in the subgroups of patients with opposite treatment responses (responding patients versus non-responders), correlation indices between scales essentially did not change. Spearman rank R were –0.872 and –0.859 in the former subgroups, –0.845 and –0.879 in the latter. The assessor was a more plausible cause of change of the KPS–ECOG PS relationship (but still a trivial cause: Spearman $R = -0.895$ for D.F. versus –0.828, for G.B.).

Two major KPS–ECOG PS equivalences have been proposed previously: one by the American Joint Committee of Cancer (AJCC) [27], the other by Minna and associates in a classic oncology textbook [1]. The two equivalence tables are reported in Table 4, along with the number and the percentage of KPS scores which could be correctly predicted using ECOG PS score (and vice versa). The equivalence by Minna and associates showed the highest level of agreement with our data: in fact, the percentage of hits was altogether 79% compared with 45% for the AJCC equivalence. Both conversion tables showed a decreasing trend of efficacy moving from the higher ECOG PS to the lower KPS scores. To improve results, a scrupulous inspection of cross-tabulated KPS and ECOG PS scores was made. A three-point conversion scale was found to exhibit the highest overall rate of success (84% rate, Table 5). However, such a simplified table was less accurate in transforming lower ECOG PS into KPS scores (61% success rate). It is interesting to note that the aforementioned three-point scale resembles quite closely the alphabetic

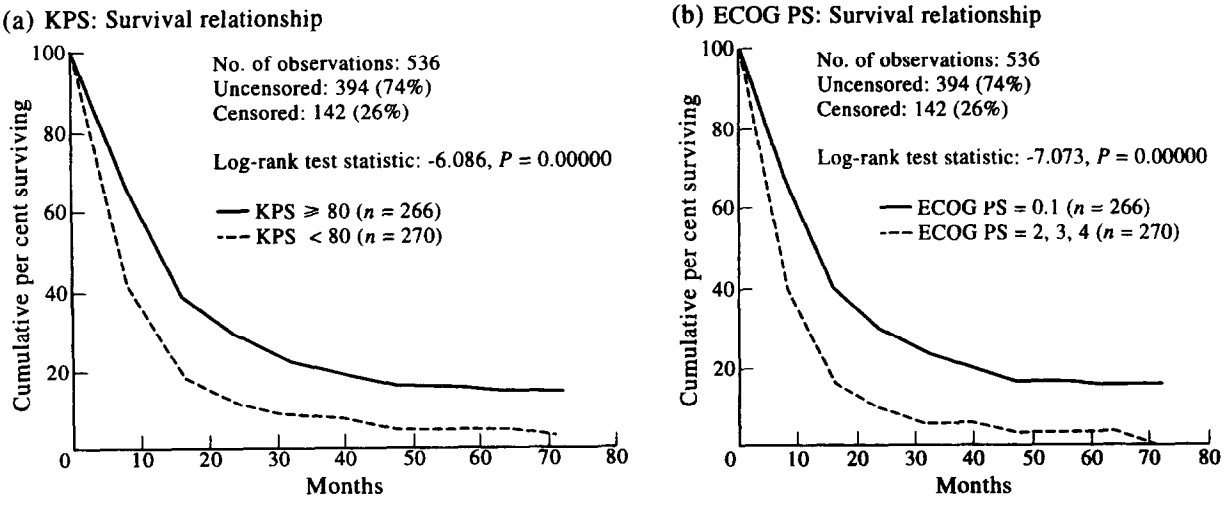

Figure 2. Survival probabilities according to (a) Karnofsky's performance status score (KPS) and (b) Eastern Cooperative Oncology Group performance status score (ECOG PS).

*Table 2. Final Cox proportional hazards model for survival**

Variable	Beta	Standard error of beta	*t* value	Exponent beta
Sex	0.4906	0.1952	2.51	1.6333
PBW	−0.0315	0.0090	−3.46	0.9690
ECOG PS	0.3415	0.0769	4.44	1.4070
TPA	0.0002	0.0001	2.30	1.0002
LDH	0.0002	0.0001	2.33	1.0001
Total white cells	−0.0001	0.0001	−2.52	0.9998
Neutrophils	0.0002	0.0001	3.16	1.0002
Clinical stage	0.1928	0.0912	2.11	1.2127
T factor	0.1480	0.0609	2.43	1.1595
N factor	0.2111	0.0596	3.54	1.2349
M factor	0.3825	0.1854	2.06	1.4660

PBW, percentage of the body weight 6 months before the diagnosis; ECOG PS, Eastern Cooperative Oncology Group performance status; TPA, tissue polypeptide antigen; LDH, lactate dehydrogenase.
* No. of valid observations: 473; $\chi^2 = 221.718$; df = 11; $P = 0.00000$.

Table 3. Cox proportional hazards regression models for survival (only variables normally used for stratification purposes)

Variable	Beta	Standard error of beta	*t* value	Exponent beta
Model A: No. of valid observations: 526, $\chi^2 = 143.808$; df = 4; $P = 0.00000$				
PBW	−0.0314	0.0083	−3.8	0.9691
Karnofsky PS	−0.0086	0.0074	−1.2	0.9914
ECOG PS	0.2044	0.1202	1.7	1.2268
Clinical stage	0.0595	0.0071	8.4	1.0613
Model B: No. of valid observations: 526, $\chi^2 = 116.246$; df = 3; $P = 0.00000$				
PBW	−0.0329	0.0082	−3.9	0.9675
Karnofsky PS	−0.0189	0.0043	−4.3	0.9812
Clinical stage	0.0601	0.0071	8.5	1.0619
Model C: No. of valid observations: 526, $\chi^2 = 156.752$; df = 3; $P = 0.00000$				
PBW	−0.0325	0.0083	−3.9	0.9679
ECOG PS	0.3166	0.0709	4.5	1.3724
Clinical stage	0.0596	0.0071	8.5	1.0614

PBW, percentage of the body weight 6 months before the diagnosis; ECOG, Eastern Cooperative Oncology Group; PS, performance status.

A, B, C classification of the same KPS [7]: grade 1: patients still able to work: KPS = 100, 90, 80 or ECOG PS = 0, 1 (work limitations of various degree, or signs and symptoms of disease may be present); grade 2: patients totally unable to work, but still able to care for themselves: KPS = 70, 60 and ECOG PS = 2 (various levels of assistance needed; however, at variance with the B classification of KPS, considerable assistance is excluded); grade 3: unable to care for self: KPS = 50 or lower and ECOG PS = 3, 4. Incidentally, the formerly described univariate analyses of survival were made comparing patients with grade 1 PS to patients with grade 2 and 3.

DISCUSSION

We started this study for one main reason. In drafting our first patient database, 6 years ago, we noted that we were unable to choose one scale of physical performance with an acceptable degree of confidence. At that time, several measures of PS had been described [2], among them the Karnofsky's scale [7] and the Eastern Cooperative Oncology Group Scale [8]. Both scales were widely used, and quoted in classic textbooks of oncology as an important part of the clinical evaluation of cancer patients [1]. In spite of the widespread acceptance, there was only limited information about their validity and reliability [10–12, 28]. Furthermore, only one study had reported a comparative evaluation of the two scales [12]. Albeit we had no doubt that performance status should be measured, we felt unable to decide which scale should be used. Hence, the decision was taken to use both scales, until the superiority of one scale could be established.

Strictly speaking, the term 'validity' refers to what may appropriately be inferred from scores on a measurement tool

Table 4. Comparison between AJCC [27] and Minna and associates [1] equivalence in the cohort studied

	AJCC [27]			Minna and associates [1]		
ECOG PS	KPS	Hits (%)*	Hits (%)†	KPS	Hits (%)*	Hits (%)†
0	100–90	122/124 (98)	122/305 (40)	100	73/124 (59)	73/79 (92)
1	80–70	426/608 (70)	426/949 (45)	90–80	528/608 (87)	528/660 (80)
2	50–60	173/686 (25)	173/366 (47)	70–60	589/686 (86)	589/755 (78)
3	30–40	14/202 (7)	14/33 (42)	50–40	111/202 (55)	111/151 (74)
4	20–10	3/36 (8)	3/3 (100)	30–20	11/36 (31)	11/11 (100)
Total		738/1656 (45)			1312/1656 (79)	

* Number and percentage of KPS scores correctly predicted by ECOG PS. † Number and percentage of ECOG PS scores correctly predicted by KPS.
AJCC, American Joint Committee of Cancer; ECOG PS, Eastern Cooperative Oncology Group performance status; KPS, Karnofsky's performance status.

Table 5. Proposed ECOG PS–KPS conversion table

ECOG PS score	KPS score	Hits (%, 95% CI)*	Hits (%, 95% CI)†
0, 1	100–80	658/732 (90, 84–96%)	658/739 (89, 82–96%)
2	70–60	589/686 (86, 78–93%)	589/755 (78, 69–87%)
3, 4	50–10	146/238 (61, 51–72%)	146/162 (90, 84–96%)
Total		1393/1656 (84, 76–92%)	

* Number and percentage of KPS scores correctly predicted by ECOG PS. † Number and percentage of ECOG PS scores correctly predicted by KPS.
ECOG PS, Eastern Cooperative Oncology Group performance status; KPS, Karnofsky's performance status; CI, confidence interval.

[13]. Validity assessment can be divided into several different subtypes of evaluation, including: face validity, content validity, construct validity, concurrent validity and predictive validity [2]. Predictive validity is an assessment of the ability of a test to predict a future outcome. Outcomes, such as response to treatment, duration of response, or survival are common in cancer literature. Predictive validity is only one line of inquiry into the validity of psychological instruments (for example, it does not provide information on 'what exactly KPS or ECOG PS measure'), but it is crucial in oncology research and practice. Oncologists take advantage of performance status measure predictive abilities, for example, when they select and stratify patients for treatment trials, predict and measure the efficacy of new drugs and therapies, and when they make prognosis assessments. As already remarked, the current study focuses on a comparison between two performance status scales (i.e. KPS and ECOG PS). Such a comparison is based on their predictive validity, or, more expressly, on their prognostic capability.

We report evidence in favour of the use of ECOG PS, which we feel is convincing. We have shown, in fact, that ECOG PS may have a better prognostic capability than KPS. Of course, either one of the scales was a strong predictor in univariate analysis and either one made a strong contribution when it was the only PS scale in multivariate analysis. However, when the setting for the analysis was identical (i.e. the same number of patients and/or the same covariates were considered), KPS showed less ability to discriminate patients with different prognosis, or it was expelled by ECOG PS from the final multivariate model started with both scales. In this study, the same patients were assessed by same assessors. It is important to emphasise that only two raters had the responsibility of PS evaluation. Estimates of the predictive validity of performance status may be somewhat imprecise in certain conditions. This would be the case if multiple raters entering patients into single trials did not apply performance status ratings exactly in the same manner. Absence of agreement between raters could create heterogeneous groups for analysis, and the estimates of predictive validity could therefore be inaccurate. Although excellent interobserver correlation has been already demonstrated for both KPS and ECOG PS [28], this problem is further reduced by lowering the number of assessors. Another point, which further strengths the evidence provided, is the large and homogeneous cohort of patients studied. We reported the survival records of all the patients seen during a 6 year period. Such patients, more than 500 in all, were affected by lung cancer (and, thus, suffered from a similar pattern of physical function impairment) and were managed similarly (our diagnostic and treatment protocols changed little during the duration of study).

A secondary issue, concerning the use of different scales of performance status, could be addressed by the design of this study. Scoring both KPS and ECOG PS, it is always possible to assess the degree of correlation between scores and to explore factors potentially affecting this correlation. Further, an equivalence between scales can be inferred. In a recent study [9], Verger and associates looked specifically at this point. The authors correlated KPS and ECOG PS, scored in 150 patients with various human cancers, and, using a linear regression analysis, derived point estimations and error prob-

abilities for a transformation. They stressed the difficulty of translating one score into another, especially in the range of lower performance status, where a wide spread of values was observed. Using a non-parametric statistic, we found a similarly high level of correlation between scales, which did not change substantially in a variety of conditions, such as pretreatment and post-treatment assessments, advanced and limited diseases, response or non-response to treatment, and different assessors. As previously reported [9], we found the equivalence by Minna and associates [1] the most suitable five-point transformation table, capable of a high number of hits (percentage of agreements = 79%). This was particularly true when transforming medium scores of ECOG PS into KPS aggregates (agreements in 59, 87, 86, 55 and 31% for, respectively ECOG PS equal to 0, 1, 2, 3 and 4), and, vice versa, when converting low and high KPS assessments into ECOG PS (percentage of hits = 92, 80, 78, 74 and 100%). However, the maximum level of agreement (84%) was obtained using a three-point conversion table: grade 1: KPS = 100, 90, 80 and ECOG PS = 0, 1 (work limitations of various degree, or signs and symptoms of disease); grade 2: KPS = 70, 60 and ECOG PS = 2 (slight or moderate levels of assistance needed); grade 3: KPS = 50 or lower and ECOG PS = 3, 4 (unable to care for self). This table is very much like the three alphabetic groups of KPS [7] and retains a sound validity, as shown by the univariate analyses of survival.

Performance status continues to be measured in patients with lung cancer, using either the KPS or the ECOG PS scale [6]. This choice is still based more on a fortuitous circumstance, including a sort of 'institutional inclination' rather than on objective elements of judgement. In contrast, we have always assessed our patients using both scales, which prevented us from developing a preference or an inclination towards a particular one. Based on our 'unbiased' findings, ECOG PS may have better predictive validity. In this sense, it should be preferred to KPS. We hope that this study might stimulate other work, possibly in patients suffering from other human malignancies, since our results may only apply to patients with lung cancer. Hopefully, other consistent reports could lead to a general consensus in the direction of a unique scale to be used. This would avoid problems of translating scales, which is never an easy task nor free from a certain margin of error.

1. Minna JD, Higgins GA, Glatstein EJ. Cancer of the lung. In De Vita VT Jr, Hellman S, Rosenberg SA, eds. *Cancer: Principles and Practice of Oncology*. Philadelphia, J.B. Lippincott Co., 1985, 507–597.
2. Orr ST, Aisner J. Performance status assessment among oncology patients: a review. *Cancer Treat Rep* 1986, **70**, 1423–1429.
3. Tummarello D, Graziano F, Isidori P, Cellerino R. Symptomatic, stage IV, non-small-cell lung cancer (NSCLC): response, toxicity, performance status change and symptom relief in patients treated with cisplatin, vinblastine and mitomycin-C. *Cancer Chemother Pharmacol* 1995, **35**, 249–253.
4. Ganz PA, Lee JJ, Siau J. Quality of life assessment. An independent prognostic variable for survival in lung cancer. *Cancer* 1991, **67**, 3131–3135.
5. Capewell S, Sudlow MF. Performance and prognosis in patients with lung cancer. *Thorax* 1990, **45**, 951–956.
6. Buccheri G, Ferrigno D. Prognostic factors in lung cancer: tables and comments. *Eur Resp J* 1994, 7, 1350–1364.
7. Karnofsky DA, Burchenal JH. The clinical evaluation of chemotherapeutic agents in cancer. In Macleod CM, ed. *Evaluation of Chemotherapeutic Agents*. New York, Columbia University Press, 1949, 199–205.
8. Zubrod CG, Scheiderman MA, Frei E, *et al*. Appraisal of methods for the study of chemotherapy in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *J Chron Dis* 1960, **11**, 7–33.
9. Verger E, Salamero M, Conill C. Can Karnofsky performance status be transformed to the Eastern Cooperative Oncology Group Scoring Scale and vice versa? *Eur J Cancer* 1992, **28A**, 1328–1330.
10. Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky performance status scale. An examination of its reliability and validity in a research setting. *Cancer* 1984, **53**, 2002–2007.
11. Coscarelli Schag C, Heinrich RL, Ganz PA. Karnofsky performance status revised: reliability, validity, and guidelines. *J Clin Oncol* 1984, **2**, 187–193.
12. Conill C, Verger E, Salamero M. Performance status assessment in cancer patients. *Cancer* 1990, **65**, 1864–1866.
13. American Psychological Association. *Standards for Educational and Psychological Tests*. Washington, DC, APA, 1974.
14. World Health Organisation. *International Histological Classification of Tumours*. Berlin, Springer, 1991.
15. Buccheri GF, Violante B, Sartoris AM, Ferrigno D, Curcio A, Vola F. Clinical value of a multiple biomarker assay in patients with bronchogenic carcinoma. *Cancer* 1986, **57**, 2389–2396.
16. Buccheri GF, Ferrigno D, Sartoris AM, Violante B, Vola F, Curcio A. Tumor markers in bronchogenic carcinoma: superiority of tissue polypeptide antigen to carcinoembryonic antigen and carbohydrate antigenic determinant 19-9. *Cancer* 1987, **60**, 42–50.
17. Buccheri GF, Biggi A, Ferrigno D, *et al*. Imaging lung cancer by scintigraphy with Indium-111 labeled $F(ab')_2$ fragments of the anticarcinoembryonic antigen monoclonal antibody FO23C5. *Cancer* 1992, **70**, 749–759.
18. Buccheri G, Biggi A, Ferrigno D, Leone A, Taviani M, Quaranta M. Anti-CEA immunoscintigraphy might be more useful than computed tomography in the preoperative thoracic evaluation of lung cancer. *Chest* 1993, **104**, 734–742.
19. UICC. *TNM Classification of Malignant Tumours*. Berlin, Springer, 1987.
20. Buccheri G, Ferrigno D, Rosso A. A Phase II study of methotrexate, doxorubicin, cyclophosphamide, and lomustine chemotherapy and lonidamine in advanced non-small cell lung cancer. *Cancer* 1993, **72**, 1564–1572.
21. Buccheri G, Ferrigno D. A randomised trial of MACC chemotherapy with or without lonidamine in advanced non-small cell lung cancer. *Eur J Cancer* 1994, **30A**, 1424–1431.
22. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981, **47**, 207–214.
23. Kaplan EL, Meier F. Non-parametric estimation from incomplete observations. *J Am Stat Assoc* 1958, **58**, 457–481.
24. Peto R, Pike MC, Armitage P. Design and analysis of randomized clinical trials requiring prolonged observation of each patient: II Analysis and examples. *Br J Cancer* 1977, **35**, 1–39.
25. Cox DR. Regression models and life tables. *J R Stat Soc* 1972, **34**, 187–220.
26. Siegel S. *Nonparametric Statistics for the Behavioural Sciences*. New York, MacGraw-Hill, 1956.
27. American Joint Committee on Cancer. Purposes and principles of staging. In Beahrs OH, Earl Henson D, Hutter RVP, Myers MH, eds. *Manual for Staging of Cancer*. Philadelphia, J.B. Lippincott Co., 1988, 3–10.
28. Sorensen JB, Klee M, Palshof T, Hansen HH. Performance status assessment in cancer patients. An inter-observer variability study. *Br J Cancer* 1993, **67**, 773–775.

Acknowledgements—The authors gratefully acknowledge the statistical advising of Ms Cinzia Brunelli. They also thank Ms Anna Cerchietti for her linguistic help.