Impact of the Treating Institution on Survival of Patients With "Poor-Prognosis" Metastatic Nonseminoma

Laurence Collette, Richard J. Sylvester, Sally P. Stenning, Sophie D. Fossa, Graham M. Mead, Ronald de Wit, Pieter H. M. de Mulder, Niels Neymark, Eric Lallemand, Stanley B. Kaye

For the European Organization for Research and Treatment of Cancer Genito-Urinary Tract Cancer Collaborative Group and the Medical Research Council Testicular Cancer Working Party

Background: **Because metastatic nonseminomatous germ cell cancer is a rare but treatable cancer, we have explored whether there is an association between the experience of the treating institution with this disease and the long-term clinical outcome of the patients, particularly patients with a poor prognosis.** ***Methods:*** **We analyzed data on 380 patients treated in one of 49 institutions participating in the European Organization for Research and Treatment of Cancer/Medical Research Council randomized trial of four cycles of bleomycin–etoposide–cisplatin followed by two cycles of etoposide–cisplatin versus three cycles of bleomycin–vincristine–cisplatin followed by three cycles of etoposide–ifosfamide–cisplatin–bleomycin, both treatment regimens given with or without filgrastim (granulocyte colony-stimulating factor). Institutions were divided into four groups based on the total number of patients entered in the trial. The groups were compared by use of the Cox proportional hazards model stratified for treatment with filgrastim and for patient prognosis as defined by the International Germ Cell Consensus Classification Group. With the use of this classification, only 65% of the patients had a poor prognosis.** ***Results:*** **Patients treated in the 26 institutions that entered fewer than five patients into the trial had an overall survival that was statistically significantly worse (two-sided** ***P*** **= .010; hazard ratio = 1.85; 95% confidence interval = 1.16–3.03) than that of patients treated in the 23 institutions that entered five patients or more. Overall survival and failure-free survival were similar among institutions that entered at least five patients. The observed effect may be related to differences in adherence to the chemotherapy protocol and in the frequency and extent of surgery for residual masses, although only the differences in dose intensity achieved statistical significance.** ***Conclusions:*** **Patients treated in institutions that entered fewer than five patients into the trial appeared to have poorer survival than those treated in institutions that entered a larger number of patients with "poor-prognosis" nonseminoma. [J Natl Cancer Inst 1999;91:839–46]**

Germ cell tumors account for only 1% of all cancers in men, but they represent the most common solid tumor in males between 15 and 34 years of age *(1)*, a population in whom all cancers are rare. Since the introduction of cisplatin-based chemotherapy in the 1970s, 70%–80% of patients with metastatic germ cell tumors can be cured *(2)*. Prognostic factor studies have led to the definition of risk groups suitable for different treatment strategies; most recently, the International Germ Cell Cancer Collaborative Group (IGCCCG) has produced a classification system *(3)* that has received broad approval. Although less toxic treatment is being investigated for patients with a good prognosis, the focus for patients with a poor prognosis has been on the use of more intensive chemotherapy. In addition to the prognostic factors with which a patient presents, survival and cure rates may also be related to the ability of the treating institution to give effective therapy *(4)* and to avoid lethal complications, factors that may vary from center to center within a country *(5–7)*. The specific experience that the treating center has with treating this type of disease may affect the outcome of an individual patient's treatment; as a consequence, future recommendations for referral to an advisory specialist unit would become necessary, particularly for patients with poor prognosis. We have explored this hypothesis by analyzing the results of a collaborative trial in patients with "poor-prognosis" germ cell tumors as a function of the treating institution.

Subjects and Methods

Patients and Treatments

Data on 380 patients treated in one of 49 institutions participating in the European Organization for Research and Treatment of Cancer (EORTC) and Medical Research Council (MRC) randomized trial 30895/TE13 were used in this analysis. From May 1990 through June 1994, patients were randomly assigned in a 2 × 2 factorial trial to receive one of four possible combinations of two treatments. For one treatment randomization, patients were randomly assigned to receive either four cycles of bleomycin–etoposide–cisplatin followed by two cycles of etoposide–cisplatin (BEP/EP) or three cycles of bleomycin–

Affiliations of authors: L. Collette, R. J. Sylvester, N. Neymark, European Organization for Research and Treatment of Cancer (EORTC) Data Center, Brussels, Belgium; S. P. Stenning, E. Lallemand, Medical Research Council (MRC) Cancer Trials Office, Cambridge, U.K.; S. D. Fossa, The Radium Hospital, Oslo, Norway; G. M. Mead, Royal South Hants Hospital, Southampton, U.K.; R. de Wit, Rotterdam Cancer Institute and University Hospital, The Netherlands; P. H. M. de Mulder, Sint Radboud University Hospital, Nijmegen, The Netherlands; S. B. Kaye, Western Infirmary, Glasgow, U.K.

Correspondence to: Laurence Collette, M.Sc., EORTC Data Center, Avenue Mounier 83/11, 1200 Brussels, Belgium (e-mail: lco@eortc.be).

See "Notes" following "References."

vincristine–cisplatin followed by three cycles of etoposide–ifosfamide–cisplatin–bleomycin (BOP/VIP-B). For the other treatment randomization, patients were randomly assigned to receive either granulocyte colony-stimulating factor (filgrastim) or nothing. This trial showed no statistically significant difference in efficacy between the two chemotherapy regimens (two-sided $P = .190$ for overall survival, two-sided $P = .214$ for time to disease progression, two-sided $P = .101$ for time to first treatment failure, and two-sided $P = .687$ for complete response rate) *(8)*. The second comparison found no difference in efficacy between patients receiving filgrastim and those not receiving filgrastim, but this comparison found that arms with filgrastim had improved compliance with treatment (two-sided $P = .031$ for more patients receiving at least six cycles of chemotherapy, two-sided $P = .001$ for a substantially reduced number of treatment delays for hematologic reasons, and two-sided $P = .001$ for fewer patients who received a dose reduction because of myelosuppression). This comparison also showed an increased dose intensity associated with the use of filgrastim *(9)*. In addition, there were fewer deaths related to toxicity in the filgrastim arm than in the no-filgrastim arms *(9)*.

End Points

The main end point for analyzing the effect of the treating institution was overall survival. Secondary end points were time to progression, failure-free survival, and rate of complete response. Overall survival was defined as the time from randomization of treatment to the death of the patient, whether due to primary cancer or another cause, or the date of the most recent follow-up for those still alive. Time to progression was defined as the time from randomization of treatment to the first occurrence of disease progression or to the date of the most recent follow-up. Response to treatment was assessed at the end of the treatment period. A complete response was registered in patients whose tumor markers had reached normal levels at the end of the chemotherapy, if residual masses were absent or completely resected and were found to contain no viable cancer cells at histologic evaluation of resected specimens. Patients with normal tumor markers and residual radiologic abnormalities but without histologic evaluation were classified as nonassessable. A partial response was identified in patients whose tumor markers had reached a plateau above the upper limit of the normal range without further decrease or in whom viable cancer cells were detected at surgery. Treatment failure included rising levels of tumor markers over a 4-week period and/or the appearance of new lesions and/or regrowth of an existing lesion (excluding growing mature teratoma or expanding cystic lesion). Failure-free survival was defined as the time from randomization of treatment to a treatment failure, a partial response to the treatment, the first reported progression of the disease, or the death of the patient, whichever occurred first. All patients who were free of all these events at their last visit to the clinic were censored at the date of the most recent follow-up.

Classification of Institutions

The definition of the 49 institutions that participated in the trial (cancer institute, university hospital, or district general hospital) was available from the EORTC Data Center or from the MRC or was obtained by mail from the principal investigator at the institution. The institutions were also classified into one of four *a priori*-defined categories according to the total number of patients entered into the trial as follows: fewer than five patients, five to nine patients, 10 to 19 patients, or 20 patients or more. The first category corresponds to one patient entered every year during the accrual period of the trial (from May 1990 through June 1994). There were 26, seven, 12, and four institutions in each category, respectively. Institutions were also classified according to their annual rate of accrual in the trial (fewer than two patients per year versus two patients or more per year). Despite the fact that some institutions started their participation much later after the trial opened, this second classification gave results very similar to those using total numbers accrued; these data, therefore, are not included.

No analysis by country was planned or performed. Such an analysis would be against the collaborative agreement between the two cancer research organizations involved in the study.

Statistical Methods

The outcome of the patients treated in the various groups of institutions was estimated by the Kaplan–Meier technique *(10)*. The Cox multivariate proportional hazards regression model *(11)* was used to compare the time to the event between the institutions. The comparisons with respect to the complete response rates were performed by use of logistic regression models *(12)*. The landmark method *(13)* was used to study the effect of chemotherapy dose intensity and of surgery (landmark = 6 months).

To correct for known differences in prognosis among the patients, the analysis was stratified for the IGCCCG prognosis group of the patient (good/intermediate versus poor/unknown) and for treatment with filgrastim. The groups with a good and intermediate IGCCCG prognosis were combined because of the small number of patients included. The 28 patients with unknown IGCCCG classification were analyzed with the poor-prognosis group after we observed the similarity between the long-term outcome in these two groups, as assessed by the Kaplan–Meier analysis. No adjustment for the chemotherapy was made (BEP/EP versus BOP/VIP-B) because we wanted to avoid having strata of very small size and because the chemotherapy regimen showed no impact on the efficacy results.

All statistical tests were two-sided, and the type I error probability was set to .05. Hazard ratios with their 95% confidence intervals (CIs) were used to summarize the results on the time-to-event end points, and odds ratios were used for presenting the results concerning the response rates.

Institutions that participated in trial TE13/30895 cross-classified by the following two criteria (*see* Table 2): the number of patients recruited in this particular trial and the overall recruitment in earlier or concurrent EORTC and MRC trials (EORTC trials 30795, 30824, 30847, 30873, 30874, and 30896; MRC trials TE04, TE05, TE09, TE10, TE11, and TE12). The cut points for this latter classification (30, 75, and 150 patients) were chosen to obtain the same marginal distribution of the institutions for the rows and columns of Table 2.

Results

The 380 patients included in this trial were treated during the period from 1990 through 1994 in one of 49 institutions throughout Europe (U.K., 27 institutions; The Netherlands, 12 institutions; Belgium, two institutions; Italy, one institution; France, one institution; Spain, one institution; Germany, one institution; Austria, one institution; Turkey, one institution; Hungary, one institution; and Norway, one institution). Overall, 55 patients (14%) were treated in one of the 26 institutions that entered fewer than five patients in this trial, 52 patients (14%) were treated in one of the seven institutions that entered five to nine patients, 174 patients (46%) were treated in one of the 12 institutions that entered 10 to 19 patients, and 99 patients (26%) were treated in one of the four institutions that each entered at least 20 patients into this trial (Table 1).

Although there is an overlap, it is evident that district general hospitals entered fewer patients than cancer institutes or university hospitals. Fourteen of 16 district general hospitals entered fewer than five patients compared with one of six cancer institutions and 11 of 27 university hospitals. In our sample, how-

Table 1. Description of the institutions*

Institution by No. of patients	No. of patients entered (%)	Type of institution†			Total No. of institutions
		No. of DGH	No. of CI	No. of UH	
<5	55 (14)	14	1	11	26
5–9	52 (14)	2	0	5	7
10–19	174 (46)	0	3	9	12
⩾20	99 (26)	0	2	2	4
Total	380 (100)	16	6	27	49

*Institutions were grouped according to the number of patients with nonseminomatous germ cell cancer entered into the trial (European Organization for Research and Treatment of Cancer and Medical Research Council randomized trial 30895/TE13).

†DGH = district general hospital; CI = cancer institute; UH = university hospital.

ever, about half of the centers that recruited fewer than five patients were district general hospitals and half were university hospitals.

The analyses by type of institution failed to detect any statistically significant influence of this factor on any of the end points. Because it was also thought that this classification was rather subjective, as a result of differences in definitions and health care policies between the European countries, this classification was dropped from further analyses.

The total recruitment by each institution in earlier or concurrent EORTC/MRC trials is shown in Table 2. A substantial amount of association between the total accrual in this particular trial (30895/TE13) and the total accrual in EORTC/MRC germ cell tumor trials from 1979 through 1995 is seen, particularly for the institutions that entered fewer than five patients: Twenty of 26 such institutions recruited 30 patients or fewer in total in the other EORTC/MRC trials, and two of them did not participate at all in these EORTC/MRC trials. This observation shows that the recruitment of the institutions in trial 30895/TE13 is probably indicative of the experience that institutions have with treating germ cell tumors, although this statement is subject to the assumption that most of the patients with germ cell tumors in those institutions were recruited into the EORTC/MRC clinical trials.

According to the current IGCCCG classification *(3)* (introduced after the study was completed), 65% of the patients meet the criteria of poor prognosis and 32% meet the criteria for the intermediate-risk category.

As far as the groups of patients being compared are concerned, no statistically significant imbalances were found between the four groups of patients with respect to baseline characteristics or treatment allocation, except for country and IGCCCG classification (Table 3; $P = .004$ and $P = .007$, respectively; χ^2 test). With two groups of institutions (fewer than five patients versus five patients or more), both imbalances lost statistical significance. However, a statistically significant trend was found with respect to the year of entry into the trial; i.e., 51% of the patients recruited by the institutions with a total accrual of one to four patients were recruited during the last 2 years of recruitment, whereas the percentage was only 34% in the institutions that entered at least five patients into this trial ($P = .046$). Nevertheless, there was no correlation between the year of entry and the classification of the patients according to the IGCCCG (35% good/intermediate risk from 1990 through 1991, 38% in 1992, and 35% from 1993 through 1994), which excludes a related difference in prognosis. At the time of this analysis, the median follow-up duration was 3 years. There was no statistically significant difference in the duration of follow-up when the institutions that entered five patients or more were compared with the institutions that entered fewer than five patients ($P = .119$; logrank test).

Overall Survival

The overall survival of the four groups was first compared by use of the unstratified Cox proportional hazards regression model. This model showed that the survival was similar in the three groups of institutions that entered at least five patients in the trial and that the survival in these three groups was statistically significantly better than the survival in the institutions that entered fewer than five patients ($P = .006$, when grouping all institutions with five patients or more; Fig. 1). The risk of death in the institutions entering fewer than five patients was estimated to be about twice that observed in any of the three groups of institutions with larger number of patients. The difference remained statistically significant after stratification for the IGCCCG risk group and for treatment with filgrastim ($P = .010$). By use of this stratified model, the risk of death in the institutions that entered fewer than five patients was estimated to be 1.85 times (95% CI = 1.16–3.03) that observed in the institutions that entered at least five patients into the trial. The 1-year survival rate estimated from the Kaplan–Meier curves was 70% (95% CI = 57%–82%) in the institutions that entered fewer than five patients and 82% (95% CI = 78%–87%) in the institutions that entered five patients or more. At 2 years, the survival rates were 62% (95% CI = 48%–75%) and 77% (95% CI = 72%–81%) in the institutions that entered fewer than five patients and at least five patients, respectively. The difference in survival rate was thus 12% after 1 year and 15% after 2 years. It is of note, however, that the difference in survival may have been augmented by the fact that, in five of the small institutions, recruitment was stopped after entry of a patient who died within 3 months, despite the date of death being more than a year before the closure of the trial.

Failure-Free Survival

The comparisons in terms of failure-free survival (Fig. 2) showed statistically significant differences between the institutions with fewer than or at least five patients in the unstratified analyses ($P = .014$) and in the stratified analyses ($P = .018$). The outcome in the three groups of institutions with high total accrual in the trial was similarly better than in the institutions that entered fewer than five patients. The risk of failure in this latter group of patients was estimated to be 1.56 times higher (95% CI = 1.09–2.27) than in the institutions that entered five or more patients. The 1-year and 2-year failure-free survival

Table 2. Classification of the institutions by total accrual in European Organization for Research and Treatment of Cancer (EORTC)/Medical Research Council (MRC) trial 30895/TE13 and by total recruitment in earlier or concurrent EORTC/MRC germ cell tumor trials

Institution by No. of patients in the 30895/TE13 trial	Total accrual in earlier or concurrent randomized MRC and EORTC trials				Total No. of institutions
	⩽30 patients	31–74 patients	75–150 patients	>150 patients	
<5	20*	2	4	0	26
5–9	3	3	1	0	7
10–19	3	2	5	2	12
⩾20	0	0	2	2	4
Total No. of institutions	26	7	12	4	49

*Two institutions entered no patient in earlier or concurrent randomized MRC/EORTC trials, and two other institutions started their participation in TE13/30895 within 6 months of its closure.

Table 3. Baseline characteristics of the patients

Characteristic	Institutional group			
	<5 patients (n = 55)	5–9 patients (n = 52)	10–19 patients (n = 174)	⩾20 patients (n = 99)
Age, y*	30 (15–52)	28 (17–65)	28 (16–57)	29 (15–58)
Days since diagnosis*	9 (0–1804)	11 (0–77)	10 (0–1008)	10 (0–480)
Country, No. of patients (%)				
U.K.	33 (60)	30 (58)	71 (41)	49 (50)
The Netherlands	13 (24)	8 (15)	65 (37)	20 (20)
Other	9 (16)	14 (27)	38 (22)	30 (30)
Site of primary tumor, No. of patients (%)				
Testis	40 (73)	44 (85)	137 (79)	79 (80)
Mediastinum	5 (9)	2 (4)	14 (8)	9 (9)
Abdomen	3 (5)	5 (10)	16 (9)	7 (7)
Other/unspecified	7 (13)	1 (2)	7 (4)	4 (4)
Lymph nodes metastases, No. of patients (%)†				
Abdominal	40 (73)	42 (81)	138 (79)	77 (78)
Mediastinal	17 (31)	11 (21)	46 (26)	29 (29)
Supraclavicular	14 (25)	10 (19)	35 (20)	14 (14)
Lung	36 (65)	29 (56)	114 (66)	60 (61)
Liver/bone/brain	14 (25)	7 (13)	42 (24)	29 (29)
Markers at entry on study, No. of patients (%)‡				
AFP >10 000 ng/mL	11 (20)	6 (12)	23 (13)	11 (11)
β-HCG >10 000 ng/mL	7 (13)	7 (13)	26 (15)	16 (16)
LDH >10× UNL	1 (2)	0 (0)	4 (2)	1 (1)
IGCCCG, No. of patients (%)§				
Good/intermediate	19 (35)	26 (50)	68 (39)	23 (23)
Poor/unknown	36 (65)	26 (50)	106 (61)	76 (77)
Treatment, No. of patients (%)‖				
BEP/EP, no filgrastim	19 (35)	16 (31)	56 (32)	33 (33)
BOP/VIP-B, no filgrastim	20 (36)	12 (23)	58 (33)	33 (33)
BEP/EP + filgrastim	7 (13)	13 (25)	31 (18)	15 (15)
BOP/VIP-B + filgrastim	9 (16)	11 (21)	29 (17)	18 (18)

*Values = median (range).

†Numbers in this section may add up to more than 100% because patients may have lymph node involvement at more than one location.

‡AFP = alpha-fetoprotein; β-HCG = human chorionic gonadotropin; LDH = lactate dehydrogenase; UNL = upper normal limit.

§IGCCCG = International Germ Cell Consensus Classification Group.

‖BEP/EP = four cycles of bleomycin–etoposide–cisplatin followed by two cycles of etoposide–cisplatin; BOP/VIP-B = three cycles of bleomycin–vincristine–cisplatin followed by three cycles of etoposide–ifosfamide–cisplatin–bleomycin.

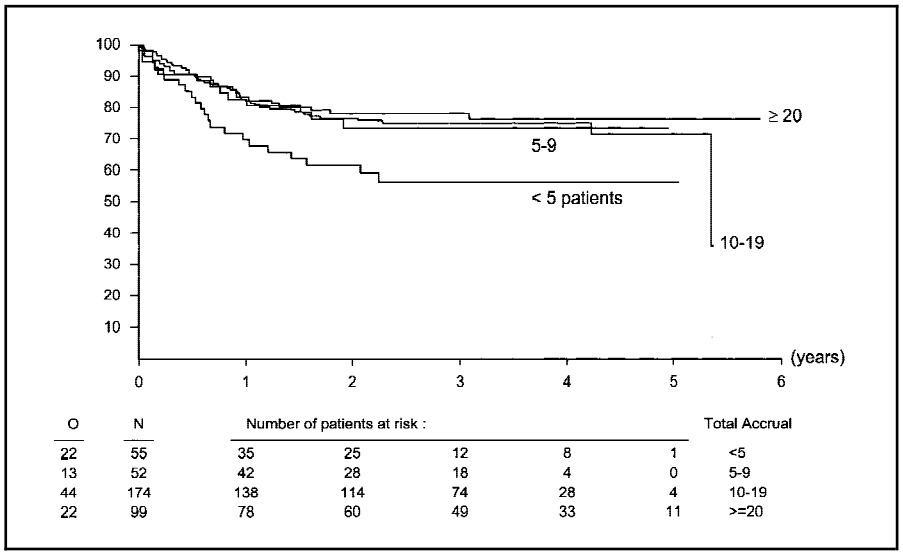

Fig. 1. Kaplan–Meier estimate of overall survival according to the total accrual of patients by the treating institution in trial 30895/TE13. O = number of deaths; N = number of patients in each group. Two-sided P = .010 in stratified analysis; hazard ratio of institutions that entered fewer than five patients versus institutions that entered five patients or more = 1.85 (95% confidence interval = 1.16–3.03). The 1-year survival rate was 70% (95% CI = 57%–82%) in the group that entered fewer than five patients and 82% (95% confidence interval = 78%–87%) in the group with at least five patients. The 2-year survival rates were 62% (95% confidence interval = 48%–75%) and 77% (95% confidence interval = 72%–81%) in the two groups of institutions, respectively.

rates for the patients treated in the institutions that entered fewer than five patients were 43% (95% CI = 29%–56%) and 38% (95% CI = 25%–51%), respectively. In the institutions that treated at least five patients in this protocol, the 1-year and the 2-year survival rates were 59% (95% CI = 54%–65%) and 55% (95% CI = 50%–61%), respectively. The difference in failure-free survival rates was 16% at 1 year and does not appear to change thereafter.

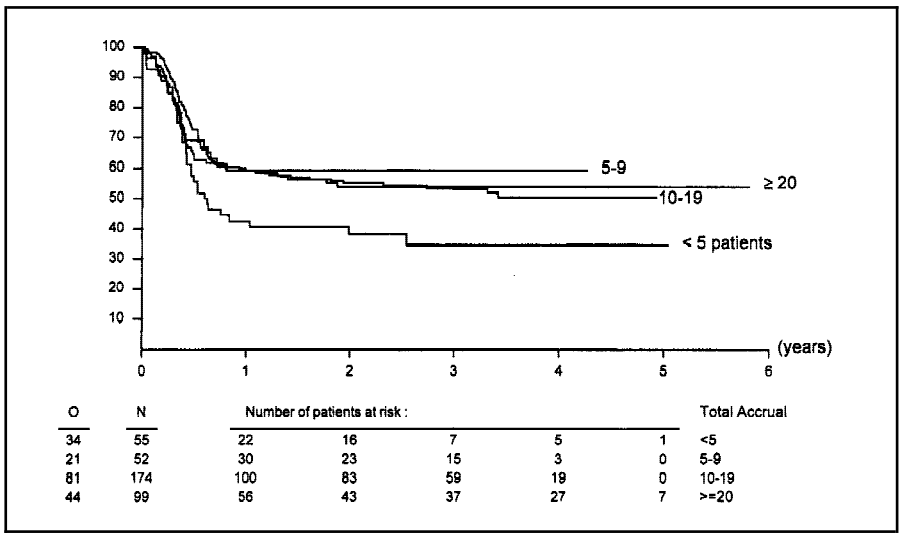

Fig. 2. Kaplan–Meier estimate of failure-free survival according to total accrual of patients by the treating institution in trial 30895/TE13. O = number of events; N = number of patients in each group. Two-sided P = .018 in stratified analysis; hazard ratio of institutions that entered fewer than five patients versus institutions that entered five patients or more = 1.56 (95% confidence interval = 1.09–2.27). The 1-year failure-free survival rate was 43% (95% confidence interval = 29%–56%) in the institutions that entered fewer than five patients and 59% (95% confidence interval = 54%–65%) in those that entered five patients or more. At 2 years, the failure-free survival rates were 38% (95% confidence interval = 25%–51%) and 55% (95% confidence interval = 50%–61%) in the two groups of institutions, respectively.

Time to Progression

A comparison of the time to disease progression confirmed these findings with statistically significantly lower progression-free rates in the institutions that entered fewer than five patients (P = .006 and P = .007 for the unstratified comparisons and stratified comparisons, respectively) and a risk of progression that was increased by a factor of 1.89 compared with the institutions that entered five patients or more (95% CI = 1.19–4.35). The 1-year progression-free rates in the two groups of institutions were 58% (95% CI = 44%–71%) and 78% (95% CI = 74%–83%), respectively. In the institutions that entered fewer than five patients, 55% (95% CI = 40%–69%) of the patients were progression free at 2 years compared with 73% (95% CI = 68%–78%) in the institutions that entered five patients or more (Fig. 3).

Complete Response

No statistically significant difference (P = .11) was observed in the rates of complete response achieved in the four groups of institutions, despite a slightly lower rate of complete response that was observed in the institutions that entered fewer than five patients (17 of 55 patients; 31%) compared with a 42% rate of complete response observed in the group of institutions that entered five patients or more (22 of 52, 73 of 174, and 43 of 99 patients in the groups of institutions that entered five to nine patients, 10 to 19 patients, and 20 patients or more, respectively). Adjustment of the model for IGCCCG classification and use of filgrastim did not change the results (P = .11).

To correct for any potential confounding effect of the country of residence in the analyses, all analyses were repeated with a stratification for the country (The Netherlands versus the U.K. versus other). The overall conclusions remained unchanged.

The above findings suggest that the patients treated in the institutions that entered fewer than five patients in this protocol have a poorer long-term outcome than those treated in the institutions that entered five patients or more. We explored potential explanations for the observed differences in prognosis by comparing the compliance with the protocol treatment and the extent of the surgery between the institutions that entered fewer than five patients and the institutions that treated at least five patients, as well as differences in treatment-induced toxicity. Other factors that might also differentiate the patient populations treated at the various hospitals, such as socio-educational factors or the performance status of the patients, could not be assessed because this information was not collected on the case report forms. (The

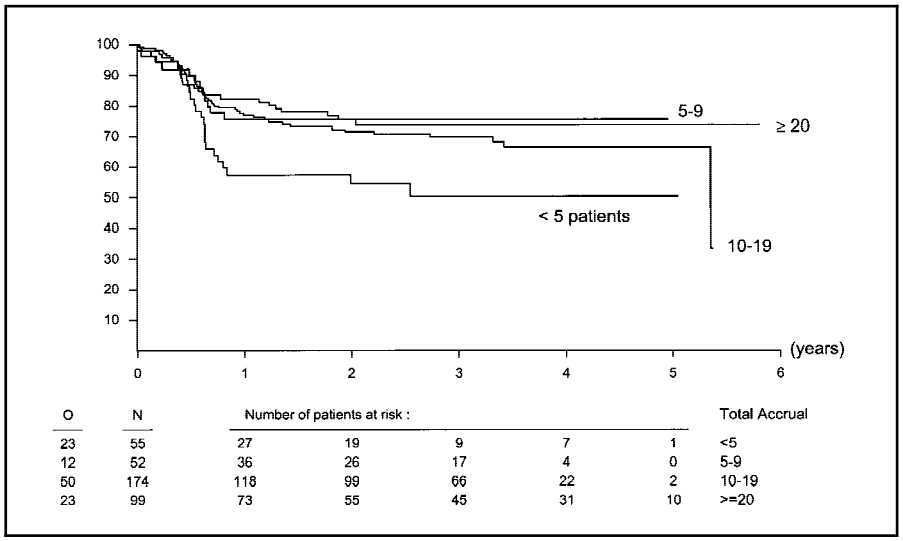

Fig. 3. Kaplan–Meier estimate of time to progression according to total accrual of patients by the treating institution in trial 30895/TE13. O = number of progressions; N = number of patients in each group. Two-sided P = .007 in stratified analysis; hazard ratio of institutions that entered fewer than five patients versus institutions that entered five patients or more = 1.89 (95% confidence interval = 1.19–4.35). The 1-year progression-free rate was 58% (95% confidence interval = 44%–71%) in the institutions that treated fewer than five patients and 78% (95% confidence interval = 74%–83%) in those that entered five patients or more. The 2-year progression-free rates were 55% (95% confidence interval = 40%–69%) and 73% (95% confidence interval = 68%–78%) in the two groups of institutions, respectively

performance status was collected only by the MRC and is not known to be a prognostic factor in this patient group.)

Compliance With the Chemotherapy Protocol

The distribution of the chemotherapy regimens BEP/EP or BOP/VIP-B was similar in the two groups of patients (Table 4). The relative dose intensity was computed as the percentage of the planned dose that was received by the patient on the basis of the number of cycles of treatment the patient actually started. Because ifosfamide and vincristine were part of the BOP/VIP-B regimen only, the comparison of the relative dose intensity of these two drugs between the two groups of patients could not be performed because of the small numbers of patients. The relative dose intensity of cisplatin and etoposide was statistically significantly lower in the institutions that treated fewer than five patients than in the institutions that treated more patients (P =

Table 4. Protocol dose adherence, surgery, and cause of death

	Institutions that entered <5 patients, No. of patients (%) (n = 55)	Institutions that entered ⩾5 patients, No. of patients (%) (n = 325)	Two-sided P
Treatment*			
BEP/EP	26 (47)	164 (51)	
BOP/VIP-B	29 (53)	161 (50)	.662†
Relative dose intensity, %‡			
Cisplatin			
BEP/EP only	97 (75–105)	99 (58–142)	.163§
BOP/VIP-B only	93 (51–106)	97 (41–161)	.015§
Both treatment arms together	95 (51–106)	98 (41–161)	.007§
Etoposide			
BEP/EP only	91 (72–103)	95 (13–104)	.419§
BOP/VIP-B only	86 (41–100)	94 (30–106)	.034§
Both treatment arms together	90 (41–103)	95 (13–106)	.036§
Six cycles or more of chemotherapy	40 (73)	256 (79)	.318†
Dose modifications			
⩾1 dose reduction	29 (53)	148 (46)	.323†
⩾1 dose delay	14 (25)	98 (30)	.480†
Hematologic toxicity‖			
WHO grade 3 or 4 neutropenia	33 (60)	165 (51)	.205†
WHO grade 3 or 4 thrombopenia	16 (29)	130 (40)	.124†
Best response to treatment			
Complete response	17 (31)	138 (42)	.107¶
Partial response	12 (22)	51 (16)	
Failure	6 (11)	30 (9)	
Early death	5 (9)	20 (6)	
Not assessable	15 (27)	86 (26)	
Residual masses after chemotherapy	42 (76)	266 (82)	.337†
Surgery, if residual masses#			
None	20 (48)	93 (35)	.093**
Biopsy	2 (5)	6 (2)	
Incomplete resection	2 (5)	25 (9)	
Complete resection	18 (43)	142 (53)	.205††
Viable cells found at histologic evaluation of resected specimen‡‡	5 (23)	21 (12)	.183§§
Cause of death			
Malignant disease	15 (27)	53 (16)	
Toxicity‖‖	7 (13)	20 (6)	.090¶¶
Other##	0 (0)	6 (2)	

*BEP/EP = four cycles of bleomycin–etoposide–cisplatin followed by two cycles of etoposide–cisplatin; BOP/VIP-B = three cycles of bleomycin–vincristine–cisplatin followed by three cycles of etoposide–ifosfamide–cisplatin–bleomycin.

†χ^2 test (1 degree of freedom).

‡Data for percent relative dose intensity are the median (range).

§Wilcoxon rank sum test.

‖WHO = World Health Organization.

¶χ^2 test on the complete response rates.

#Denominators for percentage calculations in this category are 42 and 266, respectively.

**χ^2 test comparing no surgery to any type of surgery.

††χ^2 test comparing complete resection to the rest.

‡‡Denominators for percentage calculations in this category are 22 and 173, respectively.

§§Fisher's exact test.

‖‖Treatment-induced toxicity was reported as the cause of death for five patients at institutions that entered fewer than five patients and for 14 patients at institutions that entered five patients or more. Treatment-related toxicity was reported as the cause of death for two patients at institutions that entered fewer than five patients and for six patients at institutions that entered five or more patients.

¶¶Fisher's exact test on the percent of deaths related to toxicity (treatment related plus treatment induced).

##Pulmonary embolism (two patients), lung infection (two patients), myocardial infarction (one patient), and car accident (one patient).

.007 for cisplatin and $P = .036$ for etoposide; Wilcoxon rank sum test). This difference was larger in the patients treated with the more dose-intensive BOP/VIP-B regimen than in the patients treated with BEP/EP. Because no statistically significant difference was found between the institutions with regard to the percentage of patients who received the planned six cycles of chemotherapy, the number of cycles with dose reduction, or the frequency of dose delays, the decreased dose intensity observed in the institutions that entered fewer than five patients must be related to an increased duration in the delays and to larger dose reductions. The rate of World Health Organization grade 3 or 4 hematologic toxicity was similar in both groups of patients.

Surgery

Residual masses at the end of chemotherapy were present in 42 (76%) of 55 patients treated in the institutions that entered fewer than five patients and in 266 (82%) of 325 patients treated in the institutions that entered five patients or more ($P = .337$; Table 4). Surgery for residual disease was performed in 52% of the patients treated in the institutions that entered fewer than five patients in the trial compared with 65% in other institutions ($P = .093$). The resection was macroscopically complete in 82% of the patients whose residual tumor masses were surgically removed, independent of the number of patients treated in the institution. Viable cells were found during a histologic examination of surgical specimens from five (23%) of the 22 patients whose residual tumor masses were surgically removed in the institutions that entered fewer than five patients and from 21 (12%) of the 173 patients in the institutions that entered five patients or more ($P = .183$; Fisher's exact test).

Multivariate adjustment of the analyses for chemotherapy dose intensity and surgery that used a landmark of 6 months did not change the conclusions with regard to survival and time to progression. Because more than 65% of the events for failure-free survival occurred during the first 6 months, the landmark approach could not be applied to this end point.

Toxicity

Overall, 19 patients died of treatment-induced toxicity (i.e., neutropenic sepsis, bleeding, or pulmonary fibrosis). Of these 19 patients, five were treated in an institution that entered fewer than five patients (three patients with sepsis, one patient with bleeding, and one patient with pulmonary fibrosis) and 14 were treated in an institution that entered five patients or more (10 patients with sepsis, one patient with bleeding, and three patients with pulmonary fibrosis). Another six patients died of treatment-related toxicity in the institutions that entered five patients or more; four of these patients died of postoperative complications and two of secondary acute myeloid leukemia. Two deaths due to postoperative complications were reported in the institutions that entered fewer than five patients. Thus, in total, seven (13%) of 55 patients treated in the institutions that entered fewer than five patients died of causes induced by or related to the treatment compared with 20 (6%) of 325 patients in the institutions that entered five patients or more ($P = .090$; Fisher's exact test).

Discussion

For a number of types of cancer, evidence suggests that specialist hospital care produces superior results to nonspecialist care *(14,15).* Common sense would suggest that this would apply especially to those rarer forms of cancer, such as testicular cancer, where prior experience is particularly necessary and difficult to achieve. Separate audits in Scandinavia *(6)* and Scotland *(5)* indicated a survival benefit for patients treated in specialist centers. A further study in Norway *(16)* pointed out that excellent results may also be obtained in small general oncology centers, provided that sufficient numbers of patients were treated to allow experience to accumulate within a single unit. In this study, 60% of the patients had stage I disease. Another audit in Scotland *(4)* suggested a link between treatment numbers per center and outcome, although the numbers in this study did not reach statistical significance. Retrospective audits are always open to the criticism of bias because any comparisons are inevitably not based on prospectively randomized studies and because the analysis is essentially derived from data. This situation increases the risk of making false-positive conclusions. Nevertheless, retrospective audits remain a valid tool for developing hypotheses that can later be tested in a prospective fashion.

This study involves the management of poor-prognosis testicular cancer, treatment for which may be improved by more intensive chemotherapy, better protocol dose adherence, expert surgery, and better management of treatment-induced toxicity. The management of patients with poor-prognosis cancer of the testis clearly increases the need for an appropriate supporting infrastructure of nursing and clinical care because of the increased attendant risks of treatment. Our hypothesis is that the results of this form of treatment will vary depending on the experience of the treating institution and that experience may simply be represented by the total numbers of patients recruited into the trial. Although it is conceivable that the total number of treated patients might misrepresent certain institutions (e.g., those that entered patients only in the last 1 or 2 years of the trial but had otherwise entered large numbers of patients into previous and concurrent EORTC/MRC studies), our analysis (Table 2) indicates that this is unlikely. Those centers with the worst outcome (institutions that entered fewer than five patients) also had a low rate of accrual to other studies: Twenty of the 26 institutions that recruited fewer than five patients in 30895/TE13 also recruited 30 patients or fewer for all previous or concurrent EORTC/MRC trials. Five of the institutions that entered fewer than five patients stopped recruitment when one of their first patients died shortly after entry in the trial, but four of these five institutions had recruited few or no patients in the earlier and concurrent EORTC/MRC trials. We may thus assume that these centers would have been classified as institutions that entered fewer than five patients anyway.

During recent years, MRC/EORTC trials have achieved high rates of recruitment throughout participating countries; therefore, most likely, these trends for trial recruitment accurately reflect institution experience. If recruitment is analyzed according to yearly accrual rather than according to total numbers, we conclude that those centers that entered two patients or fewer per year into this study had statistically significantly worse results than the rest of the centers. The extent of the differences is very similar to that shown for the differences according to total number recruited. Indeed, with few exceptions, the institutions that recruited two patients or fewer per year were also those that entered fewer than five patients in total.

What specific factors may underlie the observed differences? Our analysis indicates that a combination of factors may be involved in those centers that entered fewer than five patients. These factors include a greater tendency to reduce the dose of

chemotherapy, to delay treatment cycles, and to have more frequent episodes of serious treatment toxicity (including deaths from toxicity). In addition, there is a greater possibility that surgery to remove residual lesions would not be performed in these institutions. Although none of these factors in itself reaches statistical significance, the combined impact is clinically important. Indeed, in this analysis, the treating institution appears to be a prognostic factor of the same magnitude as the established pretreatment characteristics, such as marker levels and visceral metastases.

Other potentially important factors, such as the performance status or the socio-educational level of the patients, could not be assessed because this information was not requested on the case report forms of the trial.

How can these results be improved? A prospective randomized trial addressing the issue of experienced/specialist versus nonspecialist treatment center is clearly not feasible. The alternative is a prospective audit in which all patients with testicular cancer referred to each center are registered and treatment results are monitored for each risk category. Such a prospective audit would also allow information to be collected on the infrastructures of care available in each hospital. This would be useful in explaining potential observed differences in performance of treatment between the institutions. Some countries (including the U.K.) are now in the process of organizing audit systems of this type. Meanwhile, it is conceivable that referral patterns for the treatment of germ cell tumors, particularly in patients with poor prognosis, will change as the results of the present study become available to a wide audience.

In conclusion, we have shown in this trial that patients treated for poor-prognosis germ cell cancer in institutions that entered fewer than five patients in the EORTC/MRC trial 30895/TE13 have a poorer outcome than those treated in larger institutions. In this analysis, the treating institution appears to be a prognostic factor of the same magnitude as the established pretreatment characteristics. Potential explanations are related to the protocol treatment compliance and management of treatment-related toxicity. A further effect of intrinsic differences between the patient populations not accounted for in the analysis cannot be completely excluded. The trends and effects revealed in this analysis should be interpreted with caution because the comparisons are data driven. A prospective study would be needed to confirm these findings.

REFERENCES

(1) Buetow SA. Epidemiology of testicular cancer. Epidemiol Rev 1995;17: 433–49.

(2) Bosl GJ, Motzer RJ. Testicular germ cell cancer [published erratum appears in N Engl J Med 1997;337:1403]. N Engl J Med 1997;337:242–53.

(3) International Germ Cell Collaborative Group. International Germ Cell Consensus Classification: a prognostic factor-based staging system for metastatic germ cell cancers. J Clin Oncol 1997;15:594–603.

(4) Howard GC, Clarke K, Elia MH, Hutcheon AW, Kaye SB, Windsor PM, et al. A Scottish national mortality study assessing cause of death, quality of and variations in management of patients with testicular non-seminomatous germ-cell tumours. The Scottish Radiological Society and the Scottish Standing Committee of the Royal College of Radiologists. Br J Cancer 1995;72:1307–11.

(5) Harding MJ, Paul J, Gillis CR, Kaye SB. Management of malignant teratoma: does referral to a specialist unit matter? Lancet 1993;341:999–1002.

(6) Aass N, Klepp O, Cavallin-Stahl E, Dahl O, Wicklund H, Unsgaard B, et al. Prognostic factors in unselected patients with nonseminomatous metastatic testicular cancer: a multicenter experience. J Clin Oncol 1991;9: 818–26.

(7) Howard GC, Clarke K, Elia MH, Hutcheon AW, Kaye SB, Windsor PM, et al. A Scottish national audit of current patterns of management for patients with testicular non-seminomatous germ-cell tumours. The Scottish Radiological Society and the Scottish Committee of the Royal College of Radiologists. Br J Cancer 1995;72:1303–6.

(8) Kaye SB, Mead GM, Fossa S, Cullen M, de Wit R, Bodrogi I, et al. Intensive induction-sequential chemotherapy with BOP/VIP-B compared with treatment with BEP/EP for poor-prognosis metastatic nonseminomatous germ cell tumor: a randomized Medical Research Council/European Organization for Research and Treatment of Cancer study. J Clin Oncol 1998;16:692–701.

(9) Fossa SD, Kaye SB, Mead GM, Cullen M, de Wit R, Bodrogi I, et al. Filgrastim during combination chemotherapy of patients with poor-prognosis metastatic germ cell malignancy. European Organization for Research and Treatment of Cancer, Genito-Urinary Group, and the Medical Research Council Testicular Cancer Working Party, Cambridge, United Kingdom. J Clin Oncol 1998;16:716–24.

(10) Kaplan EL, Meier P: Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457–81.

(11) Cox DR. Regression models and life tables (with discussion). J R Statist Soc 1972;B34:187–202.

(12) Cox DR, Snell EJ. Analysis of binary data. 2nd ed. London (U.K.): Chapman and Hall; 1989.

(13) Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol 1983;1:710–9.

(14) Stiller CA. Centralized treatment, entry to trials and survival. Br J Cancer 1994;70:352–62.

(15) Selby P, Gillis C, Haward R. Benefits from specialised cancer care. Lancet 1996;348:313–8.

(16) Norum J, Nordoy T, Wist E. Testicular cancer treated in a minor general oncology department. Eur J Cancer 1995;31A:293–5.

NOTES

Supported by Public Health Service grants 5U10CA11488-18S2 and 2U10CA11488-19 through 2U10CA11488-28 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

Manuscript received September 14, 1998; revised March 15, 1999; accepted March 22, 1999.