Adjusting for patient selection suggests the addition of docetaxel to 5-fluorouracil–cisplatin induction therapy may offer survival benefit in squamous cell cancer of the head and neck

Jean-Pierre Pignon[a], Nathalie Syz[a], Marshall Posner[b], Robert Olivares[c], Laurence Le Lann[c], Antoine Yver[c], Ariane Dunant[a], Freddi Lewin[d], David N. Dalley[e], Adriano Paccagnella[f], Samuel G. Taylor[g], Christian Domenge[h], Jean Bourhis[i] and Madhu Mazumdar[j]

When induction chemotherapy is used in locally advanced squamous cell cancer of the head and neck (SCCHN), patients often receive cisplatin–5-fluorouracil (PF) followed by radical loco-regional therapy. Phase II studies of docetaxel–cisplatin–5-fluorouracil (TPF) induction therapy, with or without leucovorin (L), have achieved high survival rates versus those reported in phase III PF trials. However, the distribution of prognostic factors may vary between phase II and phase III study populations, making the extrapolation of phase II TPF/L results to phase III PF populations difficult. This study used a patient selection standardization method and Cox model to adjust for potential selection bias. Thus, the survival benefit from adding docetaxel into PF induction regimens in SCCHN could be more accurately assessed. The TPF/L dataset comprised 195 patients from six phase II trials. The PF dataset of 585 patients was derived from five large randomized trials included in the Meta-Analysis of Chemotherapy in Head and Neck Cancer (MACH-NC) database. TPF/L and PF datasets differed significantly concerning the distribution of several prognostic factors. Adjusting for these differences, the relative risk of death in the PF versus TPF/L datasets was 1.85 (95% confidence interval 1.37–2.49), corresponding to a 20% 2-year survival benefit ($p < 0.0001$). Sensitivity analyses confirmed that this improved 2-year survival rate of TPF/L over PF was robust, irrespective of the distribution of studied prognostic factors between treatment datasets. We conclude that this improved survival might be due either to docetaxel's pharmacologic effect or to uncontrolled prognostic factors. ***Anti-Cancer Drugs*** **15:331–340 © 2004 Lippincott Williams & Wilkins.**

Anti-Cancer Drugs 2004, **15:**331–340

Keywords: cisplatin, docetaxel, head and neck cancer, pooled analysis

[a]Unit of Biostatistics and Epidemiology, Institut Gustave-Roussy, Villejuif, France, [b]Division of Adult Oncology, Dana-Farber Cancer Institute, Boston, MA, USA, [c]Aventis Pharma, Bridgewater, NJ, USA, [d]Department of Oncology, Huddinge University Hospital, Karolinska Institute, Stockholm, Sweden, [e]Department of Medical Oncology, St Vincent's Hospital, Sydney, Australia, [f]Department of Medical Oncology, Ospedale S Giovanni e Paolo, Venezia, Italy, [g]Department of Oncology, Illinois Masonic Medical Center, Chicago, IL, USA, [h]Department of Head and Neck Surgery, Institut Gustave-Roussy, Villejuif, France, [i]Department of Radiotherapy, Institute Gustave-Roussy, Villejuif, France and [j]Division of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, NY, USA.

Sponsorship: This analysis was supported by a grant from Aventis Pharmaceuticals.

Correspondence to J.-P. Pignon, Unit of Biostatistics and Epidemiology, Institut Gustave-Roussy, 39 rue Camille Desmoulins, 94805 Villejuif Cedex, France.
Tel: +33 1 42 11 45 65; fax: +33 1 42 11 52 58;
e-mail: jppignon@igr.fr

Received 22 December 2003 **Accepted** 13 January 2004

Introduction

Most cases of squamous cell carcinomas of the head and neck (SCCHN) are locally advanced at presentation and are commonly treated with concomitant radiochemotherapy. Alternatively, induction (i.e. neoadjuvant) chemotherapy can be used. Cisplatin plus 5-fluorouracil (5-FU) (PF) represents the standard induction therapy in these patients [1,2]. However, despite relatively high response rates, the impact on overall survival is less than 5% [3,4].

The taxane docetaxel has shown significant single-agent activity in metastatic or recurrent/incurable SCCHN, producing response rates of 21–42% [5–7]. Six phase II induction studies of docetaxel–PF (TPF) with or without leucovorin (L) (TPL/L) have produced high 2-year survival rates (42–82%) and overall response rates (71–100%) in patients with locally advanced SCCHN [8–13]. Until two currently ongoing phase III studies comparing TPF-based regimens with standard PF regimens are completed, the benefit of adding docetaxel into PF regimens for SCCHN can be quantified only by comparison of the phase II TPF/L trials with 'historical controls'. One of the best controls comprises data collected from the PF arms of phase III trials.

0959-4973 © 2004 Lippincott Williams & Wilkins DOI: 10.1097/01.cad.0000124493.74538.b8

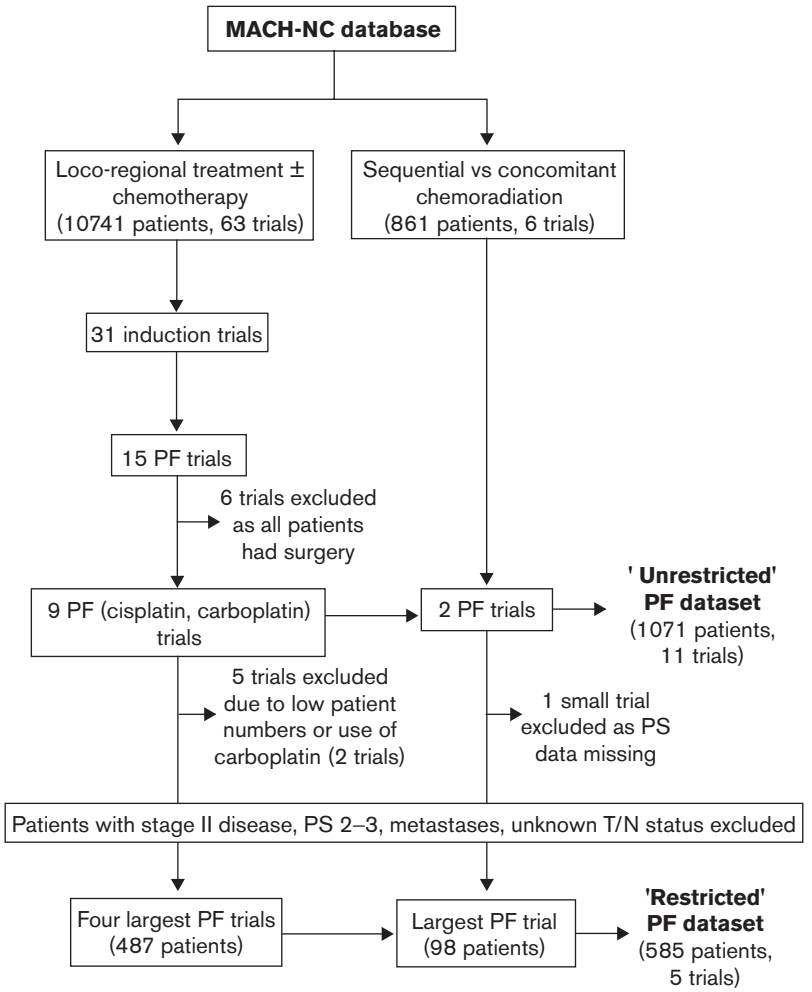
Many factors can influence treatment outcome in SCCHN [14]. These prognostic variables tend to be distributed differently in phase II study populations compared with phase III populations, making comparisons of outcomes difficult. To address this problem, Mazumdar *et al.* [15] have proposed a method (herein referred to as the 'patient selection standardization method') for adjusting the outcome obtained in phase III trials to account for the effect of patient selection in the phase II trials. This approach was applied in our analysis with the aim of anticipating the results of the ongoing phase III trials. Here, we report the results of a pooled analysis using Mazumdar's patient selection standardization methodology, as well as a classical Cox model, to compare the survival rates achieved with TPF/L in phase II trials with historical data for PF regimens after adjusting for differences in prognostic variables.

Methods

Trial identification and eligibility

At the time of analysis, six phase II trials had been conducted assessing induction TPF/L followed by radiotherapy (RT) in patients with locally advanced SCCHN [8–13]. All six studies were included in the TPF/L dataset. Comparative data on SCCHN patients treated with PF were derived from the Meta-Analysis of Chemotherapy in Head and Neck Cancer (MACH-NC) database [2]. The selection process for the two subsets of the MACH-NC data defined for the present analysis is described in Figure 1: the 'unrestricted' dataset, which included all randomized induction trials with a platinum and 5-FU, and the 'restricted' dataset, which included the largest (more than 150 patients) randomized PF induction trials. Because new data collection was needed for several of the trials, we pragmatically limited the analysis to the largest trials with cisplatin and

Fig. 1



Derivation of the 'restricted' and 'unrestricted' PF datasets from the MACH-NC database (nasopharynx carcinoma excluded) [2].

5-FU. Six trials (617 patients) in which all patients underwent surgery were excluded from both the PF datasets to make the data comparable with the TPF/L trials where only a small proportion of patients received surgery.

Patient selection and data collection

From the MACH-NC database, only those patients in the PF induction arms of studies were included in the analysis datasets. To ensure that the 'restricted' PF patients were as well matched as possible to the TPF/L patients, patients with the following characteristics were excluded: stage II SCCHN; WHO Performance Status (PS) 2–3; metastatic disease.

Data collected for each trial included: patient demographics (sex, PS, age); cancer characteristics [tumor site and disease extension (TNM classification)]; trial details [date of randomization (phase III) or treatment commencement (phase II), date of last follow-up, mortality status]; type of loco-regional treatment. Two MACH-NC database trials were updated specifically for this study (Australian and French trials).

Statistical analyses

The χ^2-test or Fisher exact test was used to compare the 'restricted' PF and the TPF/L datasets with respect to patient characteristics and loco-regional treatments. The maturity of the data and the distribution of the duration of follow-up were studied using an inverted Kaplan–Meier method [16]. The primary endpoint of the pooled analysis was survival [defined as the time from either randomization (phase III) or first drug infusion (phase II) to the date of either death from any cause or last contact] analyzed on an intent-to-treat basis using the Kaplan–Meier method [17]. As the median follow-up in the phase II studies (28 months) was lower than in the phase III studies (5 years), survival comparison was limited to 2 years of follow-up. Both the patient selection standardization method and standard survival analysis methods were used. The 'restricted' PF dataset was the primary dataset for comparison with the TPF/L dataset, with the 'unrestricted' PF dataset included to check the findings robustness.

Patient selection standardization method

The patient selection standardization method has been detailed elsewhere [15]. In our analysis, the bootstrap resampling technique was used to estimate the 2-year survival of phase III patients receiving PF therapy from the database so as to mimic the distribution in phase II studies. We adapted the method to use a combination of prognostic factors instead of a single factor. The sampling number was increased to 1000 (to improve estimate precision) and 2-year survival was estimated. The 95% confidence interval (CI) of the adjusted survival rate was then compared with the 95% CI of the unadjusted overall 2-year survival rate of the TPF/L trials.

The size of the sample used for the bootstrap—the allowable sample size (n_b)—was calculated based on the proportion of patients with each prognostic factor in the phase II studies (k_1, k_2, k_3, etc.) and the number of patients with each prognostic factor in the PF dataset (N_1, N_2, N_3, etc.). Hence, n_b was calculated as the sum of k_1% of N_1, k_2% of N_2, k_3% of N_3, etc. We then randomly sampled with replacement the 2-year survival times from the PF dataset with prognostic factors in proportions reflecting the distribution of prognostic factors in the phase II TPF/L trials. This step was repeated 1000 times to generate an empiric distribution of median adjusted 2-year survival, with its 95% CI based on the percentile method [15].

The following prognostic factors were used: age (≤ 50, 51–60, ≥ 61 years); PS (WHO 0, yes/no); tumor site (oral cavity, oropharynx, hypopharynx versus larynx versus other); tumor status [(T) T4, yes/no]; nodal status (N_{2-3}, yes/no). This selection was based on the literature, the previous study and the patient selection standardization method constraints (i.e. limitation in the number of prognostic classes). Each individual prognostic factor was considered alone, then each factor was combined with T, and lastly a combination of T, N and PS was used. Because of missing data in the database, it was not possible to study either resectability or the type of RT as prognostic factors.

Standard survival analysis methods

First, univariate analyses were performed on all the patients using a log-rank test [18,19] using the same prognostic factors as in the patient selection standardization analysis. In addition, univariate analyses of the effect of stage UICC classification (IV versus III) and country where the trial was performed (US versus other)—an indirect method for studying the effect of RT—were performed. The multivariate analyses using the Cox model [20] were then performed for prognostic factors explored using the patient selection standardization method and, in parallel, for variables with a significant prognostic value (i.e. $p < 0.05$). All statistical tests were two sided.

Results

Trial description and patient disposition

The six phase II trials assessing induction TPF/L [8–13] enrolled 198 patients (between November 1994 and May 1999), with approximately two-thirds from the US (see Table 1). Three patients with nasopharyngeal tumors and histology other than squamous cell were excluded [11], giving 195 patients for analysis. The TPF/L trials had two to five centers per trial.

The 11 trials in the 'unrestricted' PF dataset included 1071 PF-treated patients [3,4,21–29]. The five trials in the 'restricted' PF dataset were conducted between 1985 and 1993, and involved five to 18 centers per trial. Most patients were enrolled from Europe. In total, 1367 patients were enrolled, with 675 randomized to PF [3,4,22,26,28]. Of those 675 patients, 90 (13%) were excluded due to: stage II SCCHN ($n = 51$); WHO PS 2–3 ($n = 36$); metastatic disease ($n = 1$); unknown T and N ($n = 2$). Thus, the final 'restricted' PF dataset comprised 585 patients.

Patient and disease characteristics

There were statistically significant differences in patient and disease characteristics between the TPF/L and the 'restricted' PF datasets: compared with PF-treated patients, those enrolled in the TPF/L trials were younger, had a better PS and larger tumor burden, and were more often N_{2-3} (resulting in more stage IV disease) (see Table 2). Patients in the TPF/L trials tended to have fewer oral cavity and oropharynx carcinomas but more cancers of the larynx and other sites versus those in the PF trials. Nasopharynx carcinomas were reported for 9% of patients in the TPF/L trials and for 3% in the PF trials.

Treatment regimens

The treatment planned for the trials in the TPF/L and 'restricted' PF datasets is detailed in Table 1. Patients received 3–4 treatment cycles. For the TPF/L dataset, the docetaxel dose per cycle was 75–80 mg/m^2 in the three trials without leucovorin and 25–95 mg/m^2 in the three trials with leucovorin (most patients received more than 60 mg/m^2 per cycle). The cumulative dosage for 5-FU was 2800 and 3000–4000 mg/m^2 in the 5-FU-modulated and -unmodulated trials, respectively. The cisplatin dosage was 75–125 mg/m^2 in the TPF/L trials (most commonly 100 mg/m^2 or above). The PF dataset trials used similar cisplatin dosages (100 mg/m^2) but slightly higher 5-FU dosages (4000–5000 mg/m^2).

Of the six TPF/L trials, two used RT and/or surgery as planned loco-regional treatment, three used RT with neck dissection if needed, and one used RT alone. Approximately 4% of the patients had primary tumor surgery. Overall, 91% of patients received RT (hyperfractionated RT, 56%; standard RT, 23%; type of RT unknown, 12%). Among the PF trials, the GETTECneo1 trial planned only RT as loco-regional treatment [3]. In the ICC-PCP trial, surgery was performed at the investigator's discretion following chemotherapy, or for resection of residual disease after RT in the sequential chemoradiation arms [28]. In the remaining PF trials, RT and/or RT plus surgery was planned, according to the stage and the site of the disease. Overall, 22% of patients received both RT and surgery, 71% received RT alone, and 3% received surgery alone (no patients received hyperfractionated RT). The patients registered as not having loco-regional treatment per protocol were essentially patients in progression.

Survival analyses

Follow-up was truncated at 2 years in both datasets. At this point, death had occurred in 56 of 195 patients in the TPF/L dataset and 315 of 585 patients in the PF dataset. Follow-up was more complete in the PF than in the TPF/L trials: all patients from the PF dataset known to be alive had been followed for 2 or more years versus 64% of TPF/L-treated patients (1% followed for less than 6 months; 12% for 6–11 months; 7% for 12–17 months; 16% for 18–23 months).

Unadjusted survival analyses

Kaplan–Meier survival estimates for each trial in the TPF/L and the 'restricted' PF datasets are shown in Figures 2 and 3, respectively. Two-year survival estimates for the TPF/L, 'restricted' PF and unrestricted PF databases were 68% (95% CI: 60–75%), 46% (42–50%) and 52% (49–55%), respectively. The 95% CIs do not overlap, suggesting a significant survival advantage for TPF/L over PF.

Standardization of survival estimates for potential phase II design effect

Tables 2 summarizes the distribution of prognostic classes between the TPF/L and 'restricted' PF datasets. Two-year survival estimates for the 'restricted' PF dataset after adjustment using the patient selection standardization method for the distribution of survival prognostic factors in the TPF/L trials are shown in Table 3. A combined adjustment for tumor and node status, and PS yielded a 2-year survival estimate of 43% (29–57%) for PF.

A significant survival advantage for TPF/L over PF after adjusting for the same prognostic factors as in the previous analyses was also found using the Cox model: the relative risk of death in the 'restricted' PF dataset compared with the TPF/L dataset was found to be 1.80–2.31 with all 95% CI intervals excluding 1.0 (Table 3). The results of univariate and multivariate analyses on the 'restricted' PF dataset using the standard approach are shown in Table 4. In the univariate analyses, each variable (age, PS, tumor site, T and N status, stage) emerged as a significant prognostic factor. However, stage was excluded from the multidimensional analysis as it correlates closely with T and N status. In the multivariate analysis (Cox model), all the factors (except tumor site) emerged as having significant prognostic value. For the 'restricted' PF dataset, the overall relative risk of death with PF compared with TPF/L was 2.03 (1.53–2.70; $p < 0.001$) in the univariate analysis (unadjusted) and 1.85 (1.37–2.49; $p = 0.001$) in the multivariate analysis (adjusted).

Table 1 Summary of the six TPF/L trials, the five PF trials included in the 'restricted' dataset and the additional six trials (platinum plus 5-FU) included in the 'unrestricted' dataset

Trial	Country (inclusion period)	Centers	Tumor sites	Stage	Chemotherapy dose × cycle or day	Loco-regional treatment	Patients enrolled/ randomized
TPF/L dataset							
TAX017 (Posner *et al.*) [13]	The Netherlands, Belgium (1997–1998)	5	OC, OP, HP, L	III/IV	P 75–100 mg/m^2 × 4 F 3750 mg/m^2 × 4, c.i. T 75 mg/m^2 × 4	RT or S or S + RT	48
TAX708 (Posner *et al.*) [12]	US (1998–1998)	4	OC, OP, HP, L	III/IV	P 75–100 mg/m^2 × 3 F 4000 mg/m^2 × 3, c.i. T 75 mg/m^2 × 3	RT or S + RT	43
TPFL 5 (Colevas *et al.*) [8]	US (1994–1996)	2	OC, OP, HP, L, NP, O, U	III/IV	P 125 mg/m^2 × 3, c.i. F 2800 mg/m^2 × 3, c.i. L 2500 mg/m^2 × 3, c.i. T 25–60 mg/m^2 × 3[a]	RT ± S	23
TPFL 4 (Colevas *et al.*) [9]	US (1997–1998)	2	OC, OP, HP, L, NP, O, U	III/IV	P 125 mg/m^2 × 3, c.i. F 2800 mg/m^2 × 3, c.i. L 100 mg × 3, p.o. L 2000 mg/m^2 × 3, c.i. T 60 mg/m^2 × 3	RT ± S	30
OPTFL (Colevas *et al.*) [10]	US (1997–1999)	2	OC, OP, HP, L, NP, O[b], U	III/IV	P 100 mg/m^2 × 3 F 2800 mg/m^2 × 3, c.i. L 200 mg × 3, p.o. L 2000 mg/m^2 × 3, c.i. T 60–95 mg/m^2 × 3	RT ± S	34
Janinis[c] (Janinis *et al.*) [11]	Greece (1995–1997)	3	OP, HP, NP, L	III/IV	P 120 mg/m^2 × 4 F 3000 mg/m^2 × 4, c.i. T 80 mg/m^2 × 4	RT	20
'Restricted' PF dataset							
SHNG-85 (Lewin *et al.*) [26]	Scandinavia (1985–1992)	18	OC, OP, HP, L	II–IV	P 100 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT, RT + S	461
GSTTC-86 (Paccagnella *et al.*) [4]	Italy (1986–1990)	9	OC, OP, HP, O	III/IV	P 100 mg/m^2 × 4 F 5000 mg/m^2 × 4, c.i.	RT or S + RT	237
GETTECneo1 (Domenge *et al.*) [3]	France (1986–91)	6	OP	II–IV	P 100 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT	174
AHNTG (Dalley *et al.*) [22]	Australia (1986–93)	11	OC, OP, HP, NP, L, O	II–IV	P 100 mg/m^2 × 3 F 4000 mg/m^2 × 3, c.i.	RT or S or S + RT	280
ICC-PCP[d] (Taylor *et al.*) [28]	US, France (1986–1991)	≥ 5	OC, OP, HP, NP, L, O[b]	III/IV	P 100 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT, RT + S	215
'Unrestricted' PF dataset[e]							
Créteil-86 (Martin *et al.*) [27]	France (1986–1989)	>3	OC, OP, HP, L	II–IV	P 100 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT or S + RT	156
Las Palmas (Tejedor *et al.*) [29]	Spain (1987–1989)	1	OC, OP, HP, NP, L	III/IV	Cb 400 mg/m^2 × 3 UFT 14 000 mg/m^2 × 3, p.o.	RT	42
Rennes-87 (Gedouin *et al.*) [25]	France (1987–1990)	>4	OP, HP	I–IV	P 100 mg/m^2 × 3 F 4000 mg/m^2 × 3, c.i.	RT or S + RT	133
Parma (Di Blasio *et al.*) [24]	Italy (1987–1991)	>2	OC, OP, HP, L	II–IV	P 100 mg/m^2 × 3/5 F 5000 mg/m^2 × 3/5, c.i.	RT or S or S + RT	69
CFHNS (Depondt *et al.*) [23]	France (1988–1991)	9	OC, OP, HP, L	II–IV	Cb 400 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT or S + RT[f]	324
CMGH-85[d] (Adelstein *et al.*) [21]	US (1985–1988)	1	OC, OP, HP, L	II–IV	P 100 mg/m^2 × 3 F 5000 mg/m^2 × 3, c.i.	RT or RT + S	48

Study abbreviations: AHNTG: Australian Head and Neck Trial Group; CFHNS: Carboplatin French Head and Neck Study; CMGH: Cleveland Metropolitan General Hospital; GETTECneo1: Groupe d'Etude des Tumeurs de la Tête et du Cou (France); GSTTC: Gruppo di Studio sui Tumori della Testa et del Collo (Italy); ICC-PCP: Illinois Cancer Council–Paris Chicago Protocol; OPTFL: Outpatient Platin Taxotere Fluorouracil Leucovorin; SHNG: Scandinavian Head and Neck Group; TAX: Taxotere (docetaxel); TPFL: Taxotere Platin Fluorouracil Leucovorin. Site abbreviations: HP: hypopharynx; L: larynx; NP: nasopharynx; OC: oral cavity; OP: oropharynx; O: other; U: cervical nodes with unknown primary. Treatment abbreviations: Cb: carboplatin; c.i.: continuous infusion; F: 5-fluorouracil; L: leucovorin; P: cisplatin; p.o.: per os; RT: radiotherapy; S: surgery; T: docetaxel; UFT: ftorafur.

[a]25 mg/m^2 for five patients, 45 mg/m^2 for three patients, 60 mg/m^2 for the remaining patients.

[b]Patients with locally recurrent disease after surgery alone were eligible.

[c]Systematic use of granulocyte colony stimulating factor.

[d]Compared sequential with concomitant radiochemotherapy (other trials compared loco-regional treatment with loco-regional treatment plus chemotherapy).

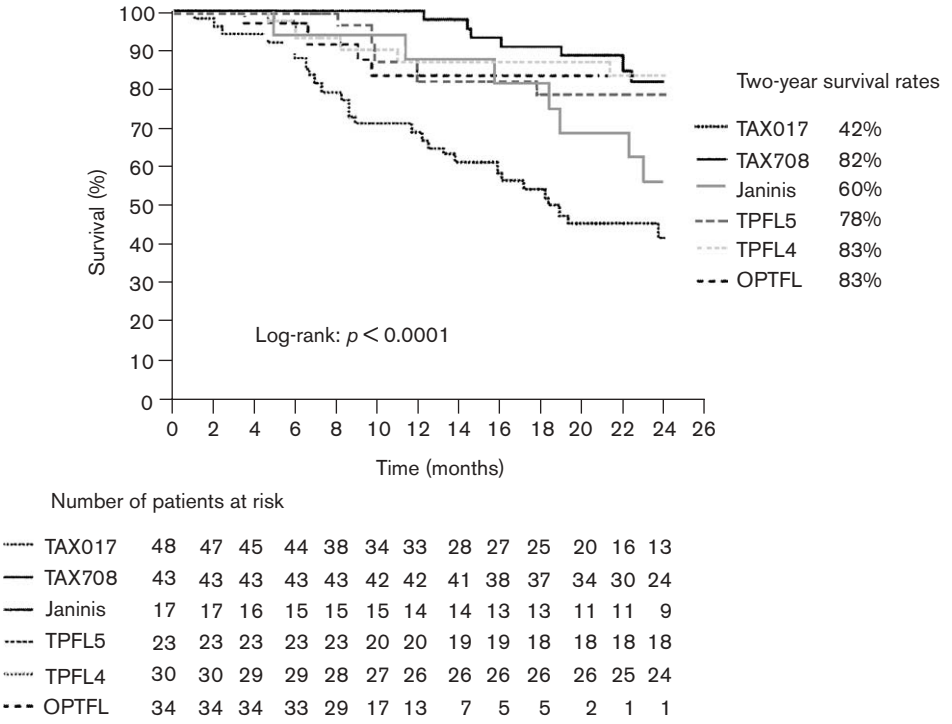
[e]Trials additional to those described under the 'restricted' PF dataset.

[f]No surgery in the case of a complete response.

Sensitivity analyses

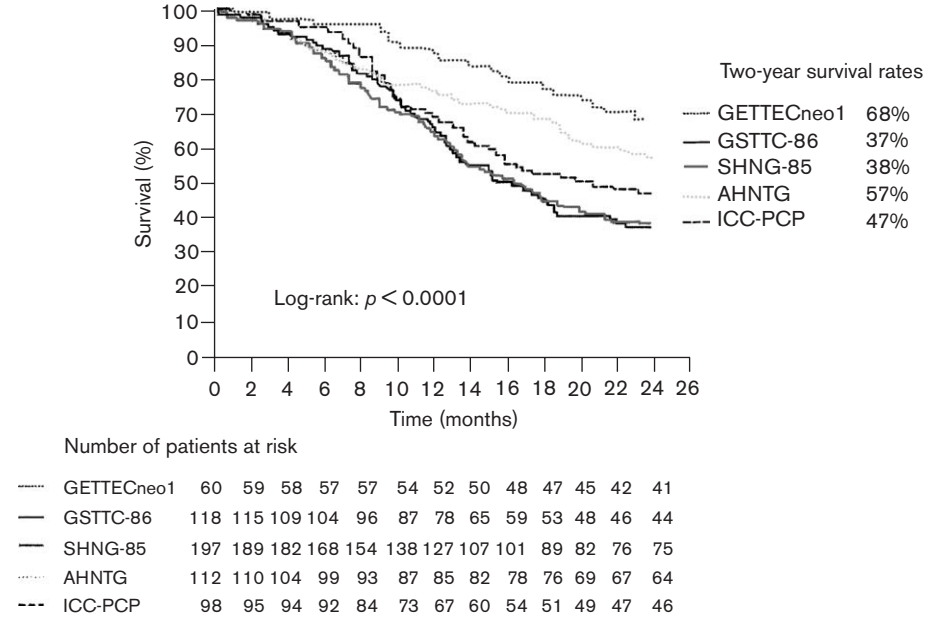
For the 'unrestricted' PF database, the unadjusted and adjusted relative risk of death with PF compared with TPF/L was 1.72. (1.31–2.27) and 1.70 (1.24–2.33), respectively. The exclusion of patients with tumors of the nasopharynx or cervical nodes with an unknown primary had little effect on the relative risk of death [1.79 (1.32–2.44)]. Sensitivity analyses were also performed after the exclusion of the OPTFL trial (not updated) and the GSTTC-86 trial (missing data on PS): exclusion of either of these trials had no significant impact on the relative risk of death [1.80 (1.32–2.46) and 1.83 (1.35–

Fig. 2



Overall survival in the TPF/L phase II trials (see Table 1 for abbreviations).

Fig. 3



Overall survival in the PF phase III trials (see Table 1 for abbreviations).

Table 2 Comparison of the distribution of patient characteristics in the TPF/L trials and the PF trials ('restricted' dataset)

Characteristic	No. (%) of patients		*p*
	TPF/L trials (*n*=195)	PF trials (*n*=585)	
Age (years)			<0.001
≤ 50	75 (39)	132 (23)	
51–60	67 (34)	192 (33)	
≥ 61	53 (27)	261 (45)	
Sex			0.382
male	157 (81)	488 (83)	
female	38 (19)	97 (17)	
WHO PS			<0.00[a]
0	92 (47)	169 (29)	
1	100 (51)	289 (49)	
2	3 (2)	0 (0)	
unknown	0 (0)	127 (22)	
T			<0.001
1	10 (5)	28 (5)	
2	27 (14)	106 (18)	
3	71 (36)	252 (43)	
4	76 (39)	198 (34)	
X, T0 or unknown	11 (6)	1 (0)	
N			<0.001
0	38 (19)	165 (28)	
1	27 (14)	170 (29)	
2	88 (45)	105 (18)	
3	40 (21)	144 (25)	
unknown	2 (1)	1 (0)	
Stage[b]			0.001
III	32 (16)	217 (37)	
IV	162 (83)	368 (63)	
unknown	1 (1)	0 (0)	
Tumor site			<0.001
oral cavity	22 (11)	138 (24)	
oropharynx	69 (36)	266 (46)	
hypopharynx	29 (15)	119 (20)	
larynx	49 (25)	42 (7)	
other	26[c] (13)	20[d] (3)	

[a]After excluding the missing values, the percentage of PS 0 and 1 was 37 and 63%, respectively, in the PF group. Among patients missing PS, six unidentified patients had a Karnofsky index <70. These patients therefore had a WHO PS ≥ 2. Thus, one can consider that 1% (6/585) of patients in the PF trials had PS=2.
[b]Based on T and N values reported for each patient, T4NX were classified stage IV.
[c]Nasopharynx (*n*=17); maxillary sinus (*n*=2); unknown primary site (*n*=7).
[d]Nasopharynx (*n*=19); unknown primary site (*n*=1).

Table 3 Estimated 2-year survival rate with PF [patient selection standardization (Mazumdar) method] and relative risk of death with PF versus TPF/L (Cox model) after adjustment of the 'restricted' PF dataset distribution of prognostic factor(s) according to the distribution in the TPF/L dataset

Prognostic factor(s) adjusted for	Mazumdar method: 2-year survival with PF (%) (95% CI) (allowable sample size)	Cox model: relative risk (95% CI) of death with PF versus TPF/L
None (unadjusted)	**46 (42–50)[a]**	**2.03 (1.53–2.70)**
TNM tumor stage: T_{1-2-3}/T_4	44 (38–49) (323)	2.09 (1.57–2.77)
Performance status: 0/1	50 (44–55) (300)	1.80 (1.35–2.39)
TNM node status: N_{0-1}/N_{2-3}	41 (36–47) (279)	2.25 (1.69–3.00)
Age: ≤ 50/51–60/ ≥ 61 years	47 (39–54) (188)	1.92 (1.44–2.55)
Site: oro-hypopharynx + oral cavity/larynx/other	49 (44–54) (337)	1.89 (1.41–2.54)
T*PS	46 (37–57) (156)	1.87 (1.40–2.49)
T*N	41 (33–49) (149)	2.31 (1.73–3.08)
T*Age	48 (33–63) (101)	1.98 (1.48–2.64)
T*Site	46 (37–55) (180)	1.99 (1.48–2.67)
T*PS*N	43 (29–57) (73)	2.15 (1.60–2.88)

[a]2-year unadjusted survival rate for TPF/L trials=68% (60–75%).

2.48), respectively]. A reduction in the relative risk of death in the PF dataset to 1.60 (1.16–2.22) was noted when the factor 'study performed in the US versus other countries' was added to the Cox model.

Discussion

In the current pooled analysis, we assessed the survival benefit of incorporating docetaxel into induction PF chemotherapy for patients with locally advanced SCCHN. In the absence of comparative phase III studies of TPF, this was achieved by comparing data from phase II trials of TPF/L with those for the induction PF arms on the MACH-NC database. MACH-NC is the largest database available on the role of chemotherapy in SCCHN. It therefore represents a valid source of 'historical data' for comparison. The largest PF trials from MACH-NC were selected for our 'restricted' PF dataset [3,4,22,26,28].

A pooled analysis of data from the six TPF/L trials generated a 2-year unadjusted survival rate of 68% (60–75%), representing a 22% absolute improvement over the unadjusted value calculated for PF [46% (42–50%)]. However, there were significant differences between the two treatment datasets with regard to prognostic factors, making simple direct comparisons of 2-year survival rates potentially inaccurate. For example, compared with PF-treated patients, patients enrolled in the TPF/L trials were younger and had better PS—factors suggesting good prognosis. However, they were more likely to be N_{2-3} or T4 combined with N_{2-3}—factors suggesting poorer prognosis. Two methods were used to adjust for the potential bias arising from the differences in patient selection: the classic Cox model [20] and a patient selection standardization method (Mazumdar) [15]. The benefits of the patient selection standardization method over the Cox model are that fewer hypotheses are needed and it does not require such large sample sizes because of the methods used to compute confidence intervals (bootstrap sampling). The limitations of this model—versus the Cox model—lie in the small number of prognostic factors that can be adjusted for and the fact that it compares timepoint survival rates instead of whole survival curves.

In the patient selection standardization analysis, TPF/L was associated with substantially better survival than PF, irrespective of the prognostic factor or combination of prognostic factors adjusted for. In all of the adjustments, with the exception of T*age, the 95% CIs for 2-year survival with PF did not overlap with those for TPF/L, indicating a statistically significant benefit of TPF/L over

Table 4 Prognostic factors for overall survival in the TPF/L and 'restricted' PF datasets and comparison of the two treatments after adjustment for prognostic factors—univariate and multivariate analyses

Characteristic	*n*	2-year overall survival (%)	Univariate analysis (*n* = 780)		Multivariate analysis (*n* = 780)	
			Relative risk (95% CI)	*p*	Relative risk (95% CI)	*p*
Trial group						
TPF/L	195	68	1	<0.0001	1	<0.0001
PF	585	46	2.03 (1.53–2.70)		1.85 (1.37–2.49)	
Age (years)						
≤ 50	207	60	1	0.002	1	0.009
51–60	259	51	1.30 (0.98–1.73)		1.24 (0.93–1.64)	
≥ 61	314	45	1.60 (1.23–2.09)		1.52 (1.15–1.99)	
WHO PS						
0	261	70	1	<0.0001	1	<0.0001
1 or unknown	519	42	2.40 (1.86–3.09)		2.04 (1.56–2.63)	
Tumor site						
oral cavity + oropharynx + hypopharynx	643	48	1.72 (1.01–2.94)	0.003	1.19 (0.69–2.05)	0.25
larynx	91	65	1.06 (0.56–1.99)		0.87 (0.46–1.66)	
other	46	66	1		1	
Tumor size						
X–0–1–2–3	506	56	1	<0.0001	1	<0.0001
4	274	42	1.55 (1.26–1.90)		1.56 (1.26–1.92)	
Nodal status						
X–0–1	403	56	1	0.002	1	<0.0001
2–3	377	46	1.38 (1.12–1.69)		1.58 (1.28–1.95)	
Stage						
3	249	60	1	<0.0001		
4	530	46	1.58 (1.26–2.00)			
unknown	1					
Country						
US	228	65	1	<0.001		
other	552	55	1.78 (1.38–2.30)			

PF and suggesting that the prognostic factors assessed were not the reason for the difference in survival rates between the two therapies. The adjusted 2-year survival rate for PF was 43–50%, and the corresponding 2-year survival advantage for TPF/L over PF was 14–26%. In the case of T*age, where the difference between the two therapies was not significant, the 95% CIs appear to have been overestimated (CIs may be overestimated by small allowable sample sizes). Results from the Cox model support these findings. According to the Cox model, the adjusted relative risk of death with PF compared with TPF/L was 1.85 (1.37–2.49) for the 'restricted' PF dataset, which translates into an absolute 2-year survival benefit for TPF/L of 20%. Even when the 'unadjusted' PF dataset was used for the comparison (a conservative estimate), a statistically significant survival advantage for TPF/L was recorded in the Cox model [adjusted relative risk of death with PF was 1.70 (1.24–2.33)]. Thus, the converging evidence of the two models suggests that the results are robust.

The robustness of these findings is further supported by the sensitivity analyses, in particular the exclusion of the 'other' tumor site category, which includes mainly nasopharynx carcinoma (which may be more sensitive to chemotherapy than other head and neck neoplasms) [30] and which led to similar results. Although the adjusted relative risk of death in the PF dataset decreased from 1.85 to 1.60 when the factor 'study performed in the US versus other countries' was added to the Cox model, the advantage of TPF/L over PF remained significant.

While every effort was made to adjust for the variations in prognostic factors between the PF and TPF/L datasets, there remain several factors we were unable to control for. This is illustrated by the persistent significant difference in survival between trials within the TPF/L dataset after adjusting for prognostic factors. This stems largely from the difference in the trial inclusion dates (1985 – 1993 for the PF phase III trials, 1994 – 1999 for the TPF/L phase II trials), and includes differences in disease staging procedures, patient selection, ancillary care and loco-regional treatment. With regard to loco-regional treatment, PF-treated patients were more likely to have undergone surgery than TPF patients, while 56% of TPF-treated patients received RT with a hyperfractionated regimen (a procedure that was not employed in the PF trials). The variable 'study performed in the US versus other countries' may partly account for the difference between the two groups since hyperfractionated RT was used in only 27% of the TPF/L European trials and none of the PF European trials. A meta-analysis of RT in head and neck cancer [31] has shown that modified fractionation (hyperfractionated and/or accelerated) RT improved 5-year survival by 3% versus standard RT. The greatest benefit was observed with hyperfractionated RT and may be as high as 8% at 5 years. Thus, the 20% absolute survival benefit of TPF over PF (after adjustment) may

be partly due to these differences in loco-regional treatment. The more frequent use of surgery, which is associated with improved prognosis, in the PF dataset is likely to have resulted in an underestimate of the survival difference between the PF and TPF/L regimens. Differences in data maturity between the two groups may also contribute to the results. With longer follow-up in the TPF/L group, the difference in survival may decrease.

In conclusion, patients in the TPF/L trials showed significantly higher survival rates than those assessed in the PF trials, irrespective of whether the prognostic factors were considered either alone or in combination. Both the patient selection standardization method of Mazumdar *et al.* [15] and the Cox model [20] suggest that the 20% difference in 2-year survival rates is not due to differences in the distribution of studied prognostic factors between the populations. The survival results may be explained by the beneficial effect of combining docetaxel with PF, by some imbalance in prognostic factors (including differences in the type of RT used) not accounted for in the analysis, or by both. The results of the ongoing randomized trials comparing TPF/L with PF should be awaited before definitive conclusions are made, as there are previous examples of well-conducted unrandomized analyses not confirmed by subsequent randomized comparisons [32]. In addition, the phase III trials will allow the effect of adding docetaxel to cisplatin–5-FU to be further defined and quantified in patients with SCCHN.

Acknowledgments

The patient selection standardization method was adapted in collaboration with Madhu Mazumdar. We thank David Adelstein, Joël Depondt, Beatrice Di Blasio, Daniel Gedouin, Michel Martin and Martín Tejedor who contributed patients to the trials under consideration, and the following statisticians who assisted in the preparation and data checking of their group trial: Gigi Pappagallo, Susan Shott, John Simes and Håkan Jonsson. We also thank Jean Pierre Armand and Catherine Hill for their expert advice on the results interpretation and manuscript preparation.

References

1 Jacobs C. Head and neck cancer in 1994: a change in the standard of care. *J Natl Cancer Inst* 1994; **86**:250–252.
2 Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. *Lancet* 2000; **355**:949–955.
3 Domenge C, Hill C, Lefebvre JL, De Raucourt D, Rhein B, Wibault P, *et al.* Randomized trial of neoadjuvant chemotherapy in oropharyngeal carcinoma. French Groupe d'Etude des Tumeurs de la Tete et du Cou (GETTEC). *Br J Cancer* 2000; **83**:1594–1598.
4 Paccagnella A, Orlando A, Marchiori C, Zorat PL, Cavaniglia G, Sileni VC, *et al.* Phase III trial of initial chemotherapy in stage III or IV head and neck cancers: a study by the Gruppo di Studio sui Tumori della Testa e del Collo. *J Natl Cancer Inst* 1994; **86**:265–272.
5 Catimel G, Verweij J, Mattijssen V, Hanauske A, Piccart M, Wanders J, *et al.* Docetaxel (Taxotere): an active drug for the treatment of patients with advanced squamous cell carcinoma of the head and neck. EORTC Early Clinical Trials Group. *Ann Oncol* 1994; **5**:533–537.
6 Couteau C, Chouaki N, Leyvraz S, Oulid-Aissa D, Lebecq A, Domenge C, *et al.* A phase II study of docetaxel in patients with metastatic squamous cell carcinoma of the head and neck. *Br J Cancer* 1999; **81**:457–462.
7 Dreyfuss AI, Clark JR, Norris CM, Rossi RM, Lucarini JW, Busse PM, *et al.* Docetaxel: an active drug for squamous cell carcinoma of the head and neck. *J Clin Oncol* 1996; **14**:1672–1678.
8 Colevas AD, Busse PM, Norris CM, Fried M, Tishler RB, Poulin M, *et al.* Induction chemotherapy with docetaxel, cisplatin, fluorouracil, and leucovorin for squamous cell carcinoma of the head and neck: a phase I/II trial. *J Clin Oncol* 1998; **16**:1331–1339.
9 Colevas AD, Norris CM, Tishler RB, Fried MP, Gomolin HI, Amrein P, *et al.* Phase II trial of docetaxel, cisplatin, fluorouracil, and leucovorin as induction for squamous cell carcinoma of the head and neck. *J Clin Oncol* 1999; **17**:3503–3511.
10 Colevas AD, Norris CM, Tishler RB, Lamb CC, Fried MP, Goguen LA, *et al.* Phase I/II trial of outpatient docetaxel, cisplatin, 5-fluorouracil, leucovorin (opTPFL) as induction for squamous cell carcinoma of the head and neck (SCCHN). *Am J Clin Oncol* 2002; **25**:153–159.
11 Janinis J, Papadakou M, Xidakis E, Boukis H, Poulis A, Panagos G, *et al.* Combination chemotherapy with docetaxel, cisplatin, and 5-fluorouracil in previously treated patients with advanced/recurrent head and neck cancer: a phase II feasibility study. *Am J Clin Oncol* 2000; **23**: 128–131.
12 Posner MR, Glisson B, Frenette G, Al-Sarraf M, Colevas AD, Norris CM, *et al.* Multicenter phase I–II trial of docetaxel, cisplatin, and fluorouracil induction chemotherapy for patients with locally advanced squamous cell cancer of the head and neck. *J Clin Oncol* 2001; **19**:1096–1104.
13 Posner MR, Vermoken JB, Janinis J, Bragas B, Yver A. Survival analysis of induction therapy with docetaxel (Taxotere) (T) containing regimens in patients with locally advanced squamous cell carcinoma of the head and neck (SCCHN). *Proc Am Soc Clin Oncol* 2001; **20**:203 (abstr).
14 Cooper JS, Farnan NC, Asbell SO, Rotman M, Marcial V, Fu KK, *et al.* Recursive partitioning analysis of 2105 patients treated in Radiation Therapy Oncology Group studies of head and neck cancer. *Cancer* 1996; **77**:1905–1911.
15 Mazumdar M, Fazzari M, Panageas KS. A standardization method to adjust for the effect of patient selection in phase II clinical trials. *Stat Med* 2001; **20**:883–892.
16 Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996; **17**:343–346.
17 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Ass* 1958; **47**:457–481.
18 Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* 1966; **50**:163–170.
19 Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer* 1977; **35**:1–39.
20 Cox DR. Regression models and life tables. *J R Statist Soc* 1972; **34**: 187–220.
21 Adelstein DJ, Sharan VM, Earle AS, Shah AC, Vlastou C, Haria CD, *et al.* Long-term follow-up of a prospective randomized trial comparing simultaneous and sequential chemoradiotherapy for squamous cell head and neck cancer. In: Salmon SE (editor): *Adjuvant Therapy of Cancer VII.* Philadelphia, PA: Lippincott; 1993, pp. 82–91.
22 Dalley D, Beller E, Aroney R, Dewar J, Page J, Phillip R, *et al.* The value of chemotherapy (CT) prior to definitive local therapy (DLT) in patients with locally advanced squamous cell carcinoma (SCC) of the head and neck (HN). *Proc Am Soc Clin Oncol* 1995; **14**:297 (abstr).
23 Depondt J, Gehanno P, Martin M, Lelievre G, Guerrier B, Peytral C, *et al.* Neoadjuvant chemotherapy with carboplatin/5–fluorouracil in head and neck cancer. *Oncology* 1993; **50**:23–27.
24 Di Blasio B, Barbieri W, Bozzetti A, Iotti C, DiSarra S, Cocconi G. A prospective randomized trial in resectable head and neck carcinoma: loco-regional treatment with and without neoadjuvant chemotherapy. *Proc Am Soc Clin Oncol* 1994; **13**:279 (abstr).
25 Gedouin D, Desprez P, Perron JJ, Fleury F, Leclech G, Miglianico L, *et al.* [Cancers of the base of the tongue and hypopharynx: results of a multicenter randomized trial of chemotherapy prior to locoregional treatment]. *Bull Cancer Radiother* 1996; **83**:104–107 [in French].

26 Lewin F, Damber L, Jonsson H, Andersson T, Berthelsen A, Biorklund A, *et al.* Neoadjuvant chemotherapy with cisplatin and 5-fluorouracil in advanced squamous cell carcinoma of the head and neck: a randomized phase III study. *Radiother Oncol* 1997; **43**:23–28.
27 Martin M, Vergnes L, Lelièvre G. A randomized study of CDDP and 5-FU as neo-adjuvant chemotherapy in head and neck cancer: an interim analysis. In: Banzet P, Holland JF, Khayat D, Weil M (editors): *Cancer Treatment: An Update*. Paris: Springer-Verlag; 1994, pp. 214–218.
28 Taylor SG 4th, Murthy AK, Vannetzel JM, Colin P, Dray M, Caldarelli DD, *et al.* Randomized comparison of neoadjuvant cisplatin and fluorouracil infusion followed by radiation vs concomitant treatment in advanced head and neck cancer. *J Clin Oncol* 1994; **12**:385–395.
29 Tejedor M, Murias A, Soria P, Aguiar J, Salinas J, Hernandez MA, *et al.* Induction chemotherapy with carboplatin and ftoratur in advanced head and neck cancer. A randomized study. *Am J Clin Oncol* 1992; **15**:417–421.
30 Fu KK. Combined-modality therapy for head and neck cancer. *Oncology* 1997; **11**:1781–1796.
31 Bourhis J, Syz N, Overgaard J, Ang KK, Dische S, Horiot J, *et al.* Conventional vs modified fractionated radiotherapy. Meta-analysis of radiotherapy in head and neck carcinoma: a meta-analysis based on individual patient data. *Int J Rad Oncol Biol Phys* 2002; **54**(**suppl**):71–72.
32 Dunn D, Babiker A, Hooker M, Darbyshire J. The dangers of inferring treatment effects from observational data: a case study in HIV infection. *Control Clin Trials* 2002; **23**:106–110.