original article

Annals of Oncology 17: 827–834, 2006
doi:10.1093/annonc/mdl033
Published online 8 March 2006

A comparative longitudinal quality of life study using the Spitzer quality of life index in a randomized multicenter phase III trial (FFCD 9102): chemoradiation followed by surgery compared with chemoradiation alone in locally advanced squamous resectable thoracic esophageal cancer

F. Bonnetain[1,2]*, O. Bouché[3], P. Michel[4], C. Mariette[5], T. Conroy[6], D. Pezet[7], B. Roullet[8], J.-F. Seitz[9], B. Paillot[10], P. Arveux[2], C. Milan[1] & L. Bedenne[1,2]

[1]Fédération Francophone de Cancérologie Digestive (FFCD) and Equipe Mixte Inserm (EMI 106), Faculté de Médecine Dijon; [2]Centre Georges Francois Leclerc, DIM, Dijon; [3]Centre Hospitalier Universitaire Robert Debré, Reims; [4]Centre Hospitalier Universitaire Charles Nicolle, Rouen; [5]Centre Hospitalier Universitaire Claude Huriez, Lille; [6]Centre Alexis Vautrin, Nancy; [7]Centre Hospitalier Universitaire Hôtel Dieu, Clermond Ferrand; [8]Centre Hospitalier Universitaire Dupuytren, Limoges; [9]Institut Paoli-Calmettes, Marseille; [10]Centre Hospitalier Universitaire du Rouen, Rouen, France

Received 7 October 2005; revised 25 January 2006; accepted 30 January 2006

Background: The aim of the study was to compare the longitudinal quality of life (QoL) between chemoradiation with or without surgery in patients with locally advanced squamous resectable esophageal cancer included in a randomized multicenter phase III trial (FFCD 9102).

Materials and methods: All patients with locally advanced resectable (T3–4 N0–1 M0) epidermoid or glandular esophageal cancer (n = 451) received induction chemoradiation. Responders (n = 259) were randomized between surgery (arm A) and continuation of chemoradiation (arm B). The Spitzer QoL Index was scored (0–10) at inclusion and at each follow-up, every 3 months during 2 years. QoL at baseline and longitudinal changes were respectively compared with univariate ANOVA and mixed-model analysis of variance for repeated measurements. The time interval between the follow-up was assessed and the same analyses were performed among survivors with 2 years of follow-up.

Results: The squamous histology was predominant in both arms. The mean QoL score decreased between baseline and the first follow-up and between the first and the second follow-ups. QoL scores at the first follow-up were comparatively worse in arm A than in arm B (7.52 versus 8.45, $P < 0.01$), whereas the longitudinal QoL study showed no difference between treatments (adjusted $P = 0.26$). Furthermore, the longitudinal QoL was not different (adjusted $P = 0.23$) among survivors with 2 years of follow-up.

Conclusions: Among patients responding to induction chemoradiation, surgery and continuation of chemoradiation had the same impact on QoL in patients with locally advanced, resectable esophageal cancer although a significantly greater decrease in the Spitzer Index was observed in the postoperative period.

Key words: quality of life, esophageal cancer, clinical trials, longitudinal, surgery, radio-chemotherapy

introduction

Esophageal cancer is a malignancy with a particularly low 5-year survival rate. The optimal management of patients with esophageal cancer and the role of surgery remain controversial [1, 2]. Until the present, surgery has been the mainstay of curative treatment in thoracic oesophageal cancer [1]. This treatment has been associated with subsequent mortality, high complication rates and quality of life deterioration [3–5].

*Correspondence to: Dr F. Bonnetain, Féderation Francophone de Cancérologie Digestive, Faculty of Medicine, 7 Boulevard Jeanne d'Arc, 21079 Dijon Cedex, France. Tel: +33-3-80-39-33-40; Fax: +33-3-80-66-82-51; E-mail: FBonnetain@dijon.fnclcc.fr

Four non-randomized studies comparing either exclusive chemoradiation and preoperative chemoradiation [6, 7] or chemoradiation alone and surgery [8, 9] showed the same survival rates between strategies with or without surgery. These results based on non-randomized trials cannot be considered as definitive. Furthermore they were influenced by tumor histology [10]. This prompted the Fédération Francophone de Cancérologie Digestive (FFCD) to set up a randomized multicenter phase III trial comparing chemoradiation (CRT) with or without surgery in locally advanced, resectable esophageal cancer. The trial was initially designed to compare overall survival and showed no survival

difference between these two arms [11]. Health-related quality of life (QoL) was a secondary end point.

Treatment selection should be based on an overall assessment of the patient's general health. From this viewpoint the QoL is now considered an important outcome in clinical cancer trials, particularly if the treatment does not result in overall survival differences [12–14]. It provides complementary and valuable information on the patient's health perception about treatment impact [15, 16]. The FDA (Food and Drug Administration) has reported that QoL is the main outcome to judge efficacy of treatment modalities when no overall survival differences are demonstrated [14, 17, 18].

Few randomized studies have prospectively assessed and compared QoL between esophageal cancer treatment modalities [4, 19, 20]. The aim of this study was to compare the longitudinal QoL between chemoradiation with or without surgery in locally advanced, resectable esophageal cancer within the framework of the randomized multicentric FFCD 9102 phase III trial.

materials and methods

eligibility criteria

Patients were eligible according to the following criteria: a locally advanced epidermoid or adenocarcinoma of the thoracic esophagus (T3–4/ N0–1/ M0); a WHO performance status of 0 to 2; eligibility for surgery (i.e. no contraindication); or a tumor judged resectable. Patients were not included if they had tracheo-bronchial involvement, if they had lost more than 15% of their body weight, or if they presented with evolutive coronary heart disease, decompensated cirrhosis or respiratory insufficiency.

Written informed consent was obtained from all patients. The protocol was approved by the Regional Ethics Committee (CCPPRB de Bourgogne, France).

study design and randomization

The trial was constructed with a two-step design. After confirming the eligibility criteria, all registered patients received similar treatment with two cycles of chemoradiation. At the end of this first step, a pre-randomization evaluation was performed. The patient was not randomized if: there was any contraindication to surgery, with the type of treatment left at the investigator's individual discretion; or a partial response was not observed and there was no improvement in dysphagia or chemoradiation was not tolerated.

Clinical response was evaluated by oesophagogram, abdominal ultrasonography, chest X-ray and, if possible, endoscopic ultrasonography. A clinical complete response was defined by the absence of dysphagia and of visible tumor on oesophagogram. A partial response corresponded to a decrease of more than 30% of the tumor length on oesophagogram, which corresponds to the WHO definition of partial response for unidimensionally measurable lesions [21] and improvement of dysphagia.

Dysphagia evaluation was part of the follow-up until death: post therapeutic period then every 3 months for 2 years and every 6 months thereafter.

Patients were randomized through a minimization program in arm A (surgery) or in arm B (continuation of chemoradiation). Stratification was carried-out according to sex, histology (epidermoid versus adenocarcinoma), differentiation (well or moderately differentiated versus poorly differentiated or undifferentiated) and response to induction treatment (complete versus partial response).

QoL

QoL was evaluated by the Spitzer QoL Index [22, 23], which is a cancer-specific QoL measurement. A score of 0 (worst) to 10 (best) was calculated after answering the five items in the areas of activity, daily life, health perceptions, social support and behavior. Each area was assessed with one item, rated on a Likert three-point scale. The Spitzer Index is a validated tool adapted to French sociological conditions and using a clinician proxy assessment [12, 23]. Due to cancer progression and/or poor health status, patients could not complete QoL questionnaires. Although self-completion was required [15, 24, 25], the Spitzer Index offered a valid QoL assessment to prevent missing data [23, 26–30].

timing of QoL assessments

The Spitzer QoL Index was scored by the clinician prior to induction chemoradiation (baseline), at the final work-up after chemoradiation CRT (arm B) or at the first post-surgical follow-up (arm A) and then at each follow-up, every 3 months during 2 years (arms A and B).

statistical analysis

All *P* values are two-sided and analyses were performed with BMDP software. According to the first objective of this study, arm A was considered equivalent to arm B if the 2-year survival rate difference was <10%. To reject this hypothesis, with a bilateral 5% type I error and a power of 80%, 360 randomized patients were required. The recruitment was stopped after the enrollment of 451 patients (259 randomized), based on the advice of an independent data monitoring committee.

Analyses were performed according to the intent-to-treat principle. The baseline characteristics of the two treatment groups were compared with a one way analysis of variance or a Student's *t*-test for continuous variables and a Pearson chi-squared test for categorical variables. Two patient subgroups according to the observed length of follow-up and survival were defined: death or drop-out <2 years versus death or drop out ≥2 years.

Depending on the treatment arm and the 2-year follow-up subgroups at each of the first five follow-ups: (1) the Spitzer Index compliance rates were described; (2) the Spitzer QoL Index and the time interval between each follow-up (in months) were described with mean, standard deviation (SD), median and range; (3) the ceiling and floor effect were evaluated with frequency [31]; and (4) univariate QoL analysis of variance was performed at baseline and at the first follow-up.

The longitudinal QoL was compared using a mixed-model analysis of variance for repeated measurements (with first order autoregressive covariance matrix) with the aim of assessing a treatment effect irrespective of the follow-up or the time interval between follow-ups [32–35]. If the follow-up effect was significant, a contrast analysis was carried out to investigate the QoL changes between: baseline and the first follow-up; the first and second follow-up; the second and third follow-up; the third and fourth follow-up; the fourth and fifth follow-up. The time interval effect was evaluated using a continuous variable representing the time (in months) between two follow-up visits. The same analyses were performed adjusted on to the 2-year follow-up variable [8] and were performed among the 2-year follow-up subgroup as exploratory analyses.

results

patients

From February 1993 to December 2000, 451 patients were registered and 444 were considered eligible. A total of 259 patients (58% of the eligible patients) were randomized: 129 were assigned to surgery and 130 to further chemoradiation continuation. The main clinical results have been presented

elsewhere [11]. The reasons for non-randomization were: no objective response or dysphagia not improved (115 patients), further chemoradiation contraindicated (32 patients who experienced grade 3–4 toxicity), surgery contraindicated (10 patients), patient's refusal (14 patients), death (eight patients) and no pre-randomization treatment (six patients).

The patient's baseline characteristics did not differ between both treatment groups (Table 1). The most common histologic type was epidermoid with a well or moderately differentiated carcinoma.

compliance with allocated arm

The compliance rate was 85% in arm A (surgery) and 97% in arm B. In arm A, 16 patients received chemoradiation and three did not receive any treatment. In arm B, one patient had surgery and three had no treatment.

therapeutic mortality

During the first 3 months after registration, 12 patients died in arm A (9.3%) and one (0.8%) in arm B ($P = 0.002$). Six-month mortality rates were 15.5% in arm A and 6.2% in arm B ($P = 0.015$).

survival

In 259 randomized patients, in arm A compared with arm B, median survival was 17.7 versus 19.3 months and 2-year survival rate was 33.6% versus 39.8%. The overall survival curves were not significantly different (relative risk of death in arm B compared with arm A = 0.88; P adjusted = 0.44) [11].

the entire study group

descriptive and univariate analyses. The Spitzer QoL Index compliance rates were greater among the patients in arm B. This varied respectively from 85% (baseline) to 66% (fifth follow-up) in arm A and from 87% to 67% in arm B (Table 2). At baseline, the QoL scores between arm A [8.44; 95% confidence interval (CI) 8.15–8.74] and arm B (8.70; CI 8.46–8.93) did not significantly differ ($P = 0.19$) (Table 2). At the first follow-up, the Spitzer QoL Index was significantly worse in the surgery arm than in arm B ($P < 0.01$): 7.52 (95% CI 6.94–8.10) versus 8.45 (95% CI 8.06–8.82). The mean difference between arms was 9% of the theoretical range score. For the entire period, the ceiling effect varied from 22% to 52% in arm A and from 32% to 46% in arm B (Table 2). In both arms, the QoL decreased between baseline and the first follow-up and between the first and second follow-up then increased until the fourth follow-up (Figure 1). Between the fourth and fifth follow-up the Spitzer Index decreased in arm B and increased in arm A (Figure 1).

Mean time interval between baseline and the first follow-up was 5.98 months (range 2–17 months) in arm A and 6.26 months (range 2–18 months) in arm B (Table 2). Thereafter, the mean time interval between the first and the second follow-up decreased (Table 2). It increased in arm A until the fifth follow-up. In arm B, it increased until the third follow-up and then decreased (Table 2).

follow-up ≥2 years versus follow-up <2 years. While at baseline the Spitzer Index compliance was better in the <2-year follow-up subgroup; thereafter the response rates were globally better among survivors with 2-year follow-up. It varied respectively from 88% (baseline) to 76% (fifth follow-up) in the 2-year survivor subgroup and 98%–63% in the other patients. As regards the 2-year follow-up subgroups, the Spitzer score did not differ at baseline (Table 2). At the first follow-up, QoL was significantly better among the 2-year subgroup ($P < 0.01$). The mean difference was equal to 10% of the theoretical range score. For the entire period, the ceiling effect varied from 31% to 59% in the 2-year and from 7% to 31% in the less than 2-year follow-up subgroups (Table 2). The mean time interval between follow-up varied from 6 to 4 months for the 2-year survivors and from 6.3 to 3.3 months for the other patients.

longitudinal QoL analysis. The Spitzer QoL Index changed significantly ($P < 0.0001$) during follow-up (Table 3); however, the treatment modality (arm A or arm B) had no influence on longitudinal QoL ($P = 0.25$). The contrast analysis between each follow-up showed a significant decrease of the Spitzer Index score between baseline and the first follow-up ($P < 0.001$) and between the first and second follow-up ($P < 0.01$).

The time interval adjusted analysis showed the same results (Table 3), i.e. there was a longitudinal change ($P < 0.0001$) and

Table 1. Demographic and clinical characteristics according to therapeutic modalities

	Arm A: surgery		Arm B: continuation of radio-chemotherapy		χ^{2*}
	N	%	*N*	%	*P* value
Gender	129		130		
Male		93.0		93.8	0.79
Female		7.0		6.2	
Response to primary treatment	129		130		
Complete		10.1		10.8	0.85
Partial		89.9		89.2	
Histology	129		130		
Epidermoid		89.1		88.5	0.86
Adenocarcinoma		10.9		11.5	
Differentiation	129		130		
Well or moderately		78.3		77.7	0.91
Poorly/undifferentiated		21.7		22.3	
Dysphagia (Atkinson)	128		130		
0/1		50.0		56.2	0.32
2/3/4		50.0		43.8	
Follow-up[a]	129		130		
<2 years		75.2		69.2	0.28
≥2 years		24.8		30.8	
At baseline	*N*	Mean (SD)	*N*	Mean (SD)	ANOVA *P* value
Age (years)	129	55.80 (10.28)	130	57.74 (10.19)	0.13
Weight (kg)	127	67.85 (13.29)	130	70.08 (12.54)	0.17

χ^{2*}, Chi-squared Pearson test significant if $P < 0.05$. ANOVA, one way analysis of variance significant if $P < 0.05$.
[a]Death or drop-out.

Table 2. Spitzer QoL Index and time interval between follow-ups description and comparison according to treatments modalities and 2-years survival

	ANOVA	Arm A: surgery					Arm B: continuation of radio-chemotherapy				
	P value	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)
Spitzer QoL index											
Baseline	0.19	110 (85.3)	8.44 (1.58)	9	1 (1.0)	10 (31.0)	113 (86.9)	8.70 (1.26)	9	4 (1.0)	10 (32.0)
First follow-up	0.01	73 (65.8)	7.52 (2.50)	8	0 (4.0)	10 (22.0)	92 (74.8)	8.45 (1.85)	9	1 (1.1)	10 (37.0)
Second follow-up		63 (69.2)	7.44 (2.61)	8	0 (1.6)	10 (23.8)	78 (71.5)	7.68 (2.99)	9	0 (6.4)	10 (37.2)
Third follow-up		46 (61.3)	7.87 (2.43)	9	0 (2.2)	10 (30.4)	56 (65.9)	8.21 (2.34)	9	0 (1.8)	10 (39.3)
Fourth follow-up		31 (57.4)	8.26 (2.67)	10	0 (3.2)	10 (51.6)	41 (64.1)	8.44 (2.40)	9	0 (2.4)	10 (46.3)
Fifth follow-up		25 (65.8)	8.76 (2.02)	10	1 (4.0)	10 (52.0)	37 (77.1)	7.81 (2.57)	9	0 (2.7)	10 (32.4)
Time interval between follow-up (months)											
0–1	0.40	111	5.98 (2.69)	5	2	17	123	6.26 (2.31)	6	2	18
1–2		91	4.11 (2.33)	4	1	11	109	3.76 (2.10)	3	0	11
2–3		75	4.21 (2.66)	3	1	15	85	4.37 (3.62)	3	0	26
3–4		54	4.67 (3.56)	3.5	0	19	64	4.26 (2.85)	3	0	12
4–5		38	5.02 (3.31)	4	1	17	48	4.29 (2.91)	3	1	14
	ANOVA	Follow-up ≥2 years					Follow-up <2 years				
	P value	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)
Spitzer QoL Index											
Baseline	0.52	64 (88)	8.67 (1.51)	9	1 (1.6)	10 (31.3)	159 (98)	8.53 (1.39)	9	3 (0.6)	10 (31.4)
First follow-up	0.002	55 (76)	8.71 (1.60)	9	4 (1.8)	10 (40.0)	110 (68)	7.70 (2.39)	8	0 (2.7)	10 (25.5)
Second follow-up		59 (82)	8.85 (1.39)	9	5 (5.1)	10 (40.7)	82 (64)	6.66 (3.21)	8	0 (7.3)	10 (24.4)
Third follow-up		52 (72)	9.02 (1.45)	9	4 (1.9)	10 (48.1)	50 (57)	7.06 (2.73)	8	0 (4.0)	10 (22.0)
Fourth follow-up		49 (72)	9.22 (1.33)	10	4 (2.0)	10 (59.2)	23 (46)	6.52 (3.34)	7	0 (8.7)	10 (26.1)
Fifth follow-up		47 (76)	8.96 (1.76)	10	1 (2.1)	10 (51.1)	15 (63)	5.80 (2.60)	6	0 (6.7)	10 (6.7)
Time interval between follow-up (months)											
0–1	0.64	72	6.01 (2.40)	6	3	18	162	6.20 (2.55)	6	2	17
1–2		72	4.42 (2.31)	4	1	11	128	3.64 (2.03)	3	0	11
2–3		72	5.42 (4.10)	4	1	26	88	3.37 (1.82)	3	0	11
3–4		68	5.26 (3.55)	4	1	19	50	3.34 (2.20)	3	0	10
4–5		62	5.35 (3.30)	4	2	17	24	2.71 (1.16)	3	1	5

Spitzer QoL Index varied from 0 (worse) to 10 (best). SD, standard deviation. One way analysis of variance significant if $P < 0.01$.
[a]Percent of responder among patients with follow-up.

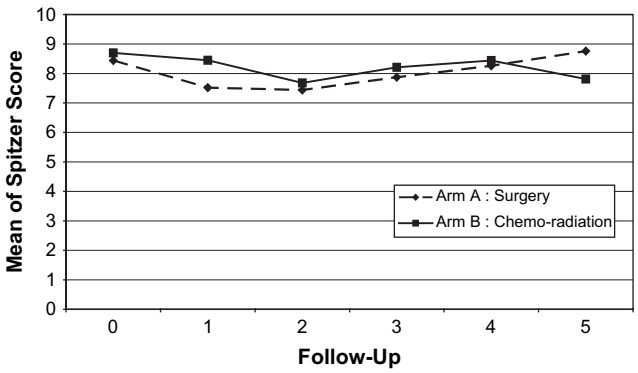

Figure 1. Mean Spitzer QoL Index at each follow-up in both arms on the whole population. Spitzer QoL index varied from 0 (worse) to 10 (best).

no treatment effect ($P = 0.26$). The time interval effect was nearly significant ($P = 0.06$): a longer time interval between follow-up was associated with a better QoL (Table 3). The contrast analysis between each follow-up revealed a significant decrease between baseline and the first follow-up ($P < 0.001$) and between the first and second follow-up ($P < 0.05$).

The 2-year follow-up status adjusted analysis confirmed that the Spitzer QoL Index changed significantly during the five follow-up visits ($P < 0.0001$) and that type of treatment had no impact on longitudinal QoL ($P = 0.40$). However, the 2-year follow-up had a significant influence on the QoL ($P < 0.0001$): whatever the treatment, at each follow-up survivor patients with 2 years follow-up had a better QoL than the patients who previously died or dropped-out. The contrast analysis showed a significant decrease of the Spitzer score between baseline and the first follow-up ($P < 0.0001$) and between the first and second follow-up ($P < 0.001$).

The time interval between follow-up adjusted analysis showed the same results (Table 3). The time interval effect was not significant ($P = 0.24$). The Spitzer Index globally changed during follow-up ($P < 0.0001$). The QoL was significantly better among the 2-year follow-up subgroup ($P < 0.0001$). The treatment arms had no influence on longitudinal QoL ($P = 0.40$). The contrast analysis confirmed a decrease of the Spitzer QoL Index between baseline and the first follow-up ($P < 0.01$) and between the first and second follow-up ($P < 0.01$).

Table 3. Analysis of longitudinal Spitzer QoL Index using a mixed model analysis of variance for repeated measurements among the whole population and the 2-year follow-up subgroup

	Among the whole population (N = 234)								Among 2-year follow-up subgroup (N = 68)			
	Model 1		Model 2		Model 3		Model 4		Model 1		Model 2	
	β[a]	P	β[a]	P	β[a]	P	β[a]	P	β[a]	P	β[a]	P
Constant	7.83	0.0000	7.56	0.0000	7.96	0.0000	7.79	0.0000	8.83	0.0000	8.83	0.0000
Follow-ups[b]		0.0000		0.0000		0.0000		0.0000		0.6747		0.6988
1	0.75	0.0002	1.01	0.0002	0.93	0.0001	1.09	0.0011	−0.19		−0.19	
2	0.14	0.0017	−0.01	0.0309	0.29	0.0004	0.20	0.0053	−0.18		−0.19	
3	−0.42	0.3690	−0.43	0.4171	−0.34	0.5973	−0.35	0.9614	−0.07		−0.07	
4	−0.24	0.4912	−0.26	0.5550	−0.23	0.9209	−0.25	0.1032	0.10		0.10	
5	−0.07	0.7527	−0.11	0.7896	−0.21	0.4009	−0.24	0.4223	0.25		0.25	
Treatment[c] Arm A versus B	−0.13	0.2533	−0.12	0.2642	−0.10	0.3573	−0.09	0.3660	−0.15	P = 0.23	−0.15	P = 0.23
Time interval between follow-ups (in months)			0.07	0.0651			0.04	0.2413			0.001	0.9704
Drop-out or death: <2 versus ≥2 years					−0.75	0.0000	−0.74	0.0000				
AIC		−1614		−1612		−1592		−1592		−556		−556

[a]Regressions terms.
[b]Contrast analysis: 1, baseline versus first follow-up; 2, first versus second follow-up; 3, second versus third follow-up; 4, third versus fourth follow-up; 5, fourth versus fifth follow-up.
[c]Treatment: arm A (surgery) versus arm B (continuation of RCT).
AIC, Akaike's information criterion.

the 2-year follow-up subgroup

descriptive and univariate analysis. In the 2-year follow-up subgroup (n = 68), the Spitzer QoL Index response rates were better among the patients in arm B. It varied from 91% (baseline) to 67% (fifth follow-up) in arm A and from 87% to 83% in arm B. The QoL at baseline did not differ between arms A and B (Table 4). At the first follow-up, the Spitzer score in arm B (9.21; 95% CI 8.84–9.58) was significantly better than in arm A (7.95; 95% CI 7.08–8.82; P <0.01). The mean difference between arms was 1.26 points, i.e. approximately 13% of the theoretical range score.

According to baseline, the QoL decreased in arm A and increased in arm B at the first follow-up (Figure 2). Then, it increased in arm A until the fourth follow-up. The Spitzer Index decreased in arm B at the second follow-up and then remained constant (Figure 2). The time interval between baseline and first follow-up did not differ between treatment modalities (Table 4). In both arms, it was shorter between the first and the second follow-up and longer afterwards (Table 4).

longitudinal QoL analysis. The mixed-model analysis of variance for repeated measurements, whether or not adjusted for the time interval between follow-ups (Table 3) revealed that the longitudinal QoL did not differ between arms (P = 0.23). The Spitzer Index score did not change significantly during the follow-up (P = 0.70). The time interval between follow-ups had no influence on the longitudinal QoL (P = 0.97).

discussion

Our study demonstrated that in patients responding to induction chemoradiation and with a predominant squamous histology, surgery or continuation of chemoradiation had the same overall impact in locally advanced, resectable esophageal cancer. There was no difference in the overall survival between arms [11]. The longitudinal QoL measured by the Spitzer Index did not differ between these treatment modalities. In the chemoradiation arm, the 3-month mortality was lower, the hospital stay was shorter than in the surgery arm, but there was more frequent palliative procedures against dysphagia (stent or dilatation) (P <0.0001) [11]. Nevertheless, this did not appear to influence longitudinal QoL in patients, particularly among long-term survivors.

As previously shown, the longitudinal QoL of the 2-year survivors (without drop-out) was significantly better than other patients in both arms [3, 20, 36]. Patients who die within 2 years often suffer from an irreversible degradation of QoL whatever the treatment [3]. However, our study revealed that among the latter survivors, QoL appeared to remain constant during the trial and suggested an association between overall survival and QoL level whatever the treatment [36, 37]. The poor long-term survival also implied that few patients had a long-term QoL follow-up, reducing the power of our longitudinal study.

Despite these overall results, there was a greater QoL decrease at the first follow-up in the surgery arm and a better QoL was observed in the chemoradiation arm. Some studies showed the same impact among the 2-year survivors [3, 20], whereas another study had shown an unaltered global evaluation of QoL, although physical symptoms increased at postoperative assessment [38]. In our study, among survivors with at least 2 years of follow-up, the mean difference between arms represented 13% of the Spitzer theoretical range score. This difference was greater than 10% and was considered clinically significant [12, 39]. In the entire population, the difference was

Table 4. Description and comparison of Spitzer QoL Index and time interval between follow-ups according to treatment modalities among the 2-year follow-up subgroup

	ANOVA	Arm A: surgery					Arm B: continuation of radio-chemotherapy				
	P value	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)	*N* (%)[a]	Mean (SD)	Median	Min (%)	Max (%)
Spitzer QoL Index											
Baseline	0.14	29 (91)	8.35 (2.00)	9	1 (3.4)	10 (34.5)	35 (87)	8.94 (0.87)	9	7 (5.7)	10 (28.6)
First follow-up	0.01	22 (69)	7.95 (1.93)	9	4 (4.5)	10 (27.3)	33 (82)	9.21 (1.05)	9	5 (3.0)	10 (48.5)
Second follow-up		25 (78)	8.68 (1.46)	9	5 (4.0)	10 (36.0)	34 (85)	8.97 (1.34)	9	5 (5.9)	10 (44.1)
Third follow-up		22 (69)	9.04 (1.59)	9.5	4 (4.5)	10 (50.0)	30 (75)	9.00 (1.36)	9	5 (6.7)	10 (46.7)
Fourth follow-up		19 (63)	9.26 (1.63)	10	4 (5.3)	10 (73.7)	30 (79)	9.20 (1.13)	9.5	5 (3.3)	10 (50.0)
Fifth follow-up		18 (67)	9.17 (2.12)	10	1 (5.6)	10 (66.7)	29 (83)	8.83 (1.51)	9	4 (3.4)	10 (41.4)
Time interval between follow-up (months)											
0–1	0.66	32	5.87 (2.18)	5.5	3	11	40	6.12 (2.56)	6	3	18
1–2		32	4.53 (2.35)	4	1	11	40	4.32 (2.29)	3	1	11
2–3		32	5.12 (3.19)	4	1	15	40	5.65 (4.68)	4	1	26
3–4		30	6.20 (3.99)	6	2	19	38	4.53 (3.02)	3	1	12
4–5		27	5.85 (3.58)	5	2	17	35	4.97 (3.06)	3	2	14

Spitzer QoL Index varied from 0 (worse) to 10 (best). SD, standard deviation.
ANOVA, one way analysis of variance significant if $P < 0.01$.
[a]Percent of responder among patients with follow-up.

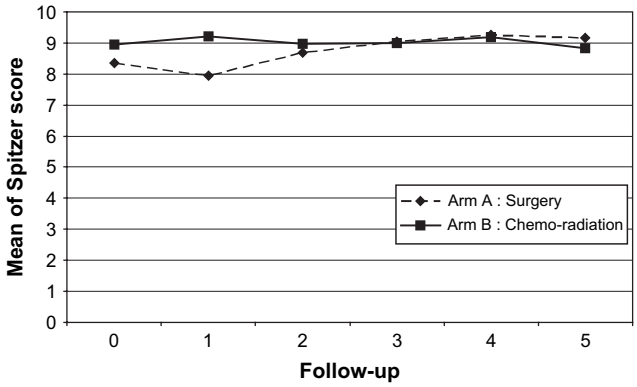

Figure 2. Mean Spitzer QoL Index at each follow-up according to treatment modalities among the 2-year follow-up subgroup.

9% and the Spitzer score decreased at the first and second follow-up in both arms. Regarding QoL, the surgery and the post-surgery experience seemed initially to be poor while there was no improvement of survival. Although the longitudinal QoL did not differ between arms, these results should be considered in therapeutic decision. We have not explored the QoL of non-responder patients. As demonstrated by a German study [40], response to induction chemotherapy was identified as a single independent prognostic factor (HR = 0.30; 95% CI 0.19–0.47). This study confirmed our survival results in responder patients. Regarding our results among responder patients, surgery should not be considered as standard treatment. However, among non-responder patients, surgery seemed to improve survival suggesting that non-responder to induction chemoradiation could have a potential benefit of curative surgery [40]. This hypothesis should be proven by larger trials and/or meta-analyses.

In this study, an analysis was performed in the subgroup of patients who lived longer than 2 years with follow-up. It was carried-out with the aim of reducing the biases of missing values and drop-out patterns [27, 41–43], the potential confounding effect of recurrences and the time interval variability between follow-ups [35]. This approach only partially reduces these biases [41]. However, based on the QoL study, it represented a more specific health and care profile and it is reasonable to assume that the missing data mechanism is homogeneous within this defined health state trajectory and trial participation [3, 20, 44]. Nevertheless, the results of the multivariate analysis of variance for repeated measurements revealed some interesting points. In the entire population, the time interval between follow-ups effect was nearly significant: the longer the time interval, the better the QoL. In contrast, the 2-year follow-up adjusted analysis showed that the time interval did not influence the Spitzer Index score. Therefore, the time interval between follow-ups could be associated with the 2-year survival: it was consistently shorter for those who died or dropped-out within 2 years. Among survivors with 2-year follow-up, the time interval had no direct influence on QoL assessment, whatever the treatment. As sensitivity analyses, the time interval adjusted analyses confirmed our results with robustness [44]. Further statistical models and analysis taking into account this potential bias are warranted, for example, time until definitive QoL deterioration should be an alternative [45].

The Spitzer Index is a global QoL index and has been assessed by clinicians so it could be considered as not relevant to capture the perceived QoL [15, 25]. Physician or other observer ratings of patient's health-related QoL have previously obtained a moderate correlation with self-appraised QoL [25]. Physicians seemed to underestimate QoL (by 10% on average) and the severity of symptoms [12, 24]. In contrast, professionals can also provide useful information particularly regarding the more concrete, observable aspects of QoL [46–48]. Specific instruments to assess esophageal cancer have been shown to be more accurate and responsive to evaluate QoL [49–51].

However, a global QoL index remains relevant to compare the global impact of treatment arms during times [12, 29, 30]. Furthermore, the use of a clinician proxy to assess QoL could improve data collection when patients are unable to complete a questionnaire [46–48]. We were not able to maintain optimal compliance rates during the study. According to non-random missing data, the QoL in both arms was probably over estimated in the last follow-ups [42]. The acceptability of QoL assessment should be specifically evaluated and optimized in further FFCD trials to prevent missing data [26, 27].

The longitudinal change may have been compromised by ceiling effects. Therefore, the Spitzer Index cannot be used to detect any specific difference in change between arms at a higher level. These psychometric properties were one of the principal limitations of our results concerning the positive impact of the treatment modalities, particularly in the 2-year follow-up subgroup. In contrast, agreement between patients and proxies seemed to be greater when QoL is very poor or very good, known respectively as the floor and ceiling effect [24]. Therefore, the global positive or negative impact of the treatment modalities on QoL was not compromised but could be underestimated.

To our knowledge and despite other studies, no other randomized study comparing chemoradiation with or without surgery in locally advanced, resectable esophageal cancer has investigated longitudinal QoL [1, 2, 40]. Considering the limits of our study and the predominance of squamous histology, surgery and exclusive concomitant chemoradiation had the same impact on the longitudinal Spitzer QoL Index. Further randomized studies among responder and non-responder patients to induction chemoradiotherapy with multidimensional and esophageal cancer specific QoL instruments are warranted to confirm our results.

acknowledgements

This study was supported by grants from the Ligue Nationale Contre le Cancer (LNCC), the Fonds de la Recherche de la Société Nationale Française de Gastroentérologie (SNFGE), the Programme Hospitalier pour la Recherche Clinique (PHRC) and the Association pour la Recherche contre le Cancer (ARC).

Presented in part at the Thirty-ninth Annual Meeting of the American Society of Clinical Oncology, 2003.

The authors thank Richard Medeiros, Rouen University Hospital Medical Editor, for his valuable contribution in editing the manuscript. We are indebted to Mrs. Cécile Girault for data management.

In addition to the authors, the following investigators participated in the FFCD 9102 trial: J.-P. Herr (Chalon sur Saône Private Hospital Saint Marie); J. Butel, P. Desselle, J. C. Brice (Abbeville Hospital); B. Tissot (Agen Hospital); A. Votte-Lambert, J. P. Joly (Amiens University Hospital); P. Burtin, J. P. Arnaud (Angers University Hospital); P. Cellier (Angers Anticancer Centre); F. Estermann (Annemasse Hospital; B. Chauvet (Ste Catherine Institute Avignon); E. Maringe (Beaune Hospital); F. Ozanne (Beauvais Hospital); F. Varlet (Béthune Private Hospital Ambroise Paré); Y. Becouarn, A. Avril (Bordeaux Anticancer Centre Bergonié); P. Rougier, B. Nordlinger (Boulogne-Billancourt University Hospital Ambroise Paré); M. Vincendet (Boulogne Sur Mer Private Hospital); J. Charneau (Boulogne Sur Mer Hospital); D. Pillon (Bourg En Bresse Hospital); N. Stremsdoerfer, M. Pelletier (Bourgoin-Jallieu Hospital); M. C. Clavero-Fabri (Bligny Hospital—Briis Sous Forges); B. Leduc (Brive Hospital); P. Segol, L. P. Argouach, A. Roussel, J. Maurel, R. Salame (Caen University Hospital); J. Lacourt, P. Janoray, O. Ruget (Chalon Sur Saône Private Hospital Ste Marie); D. Baudet-Klepping (Chalon Sur Saône Hospital); G. Dupont (Chaumont Private Hospital); G. Bommelaer (Clermont-Ferrand University Hospital); P. Ruszniewski, P. Hammel (Beaujon University Hospital, Clichy); S. Chaussade, B. Dousset (Cochin University Hospital); B. Denis, J. D. Wagner, E. Tamby, T. Petit, A. M. Weiss (Colmar Hospital Pasteur); J. C. Barbare (Compiègne Hospital); J. L. Jouve, J. M. Phelip, P. Senesse, C. Michiels (Dijon University Hospital); P. Maingon, B. Coudert, J. Fraisse (Dijon Anticancer Centre Georges-François Leclerc); L. Gasnault, J. H. Gstach, B. Guichard (Dunkerque Hospital); M. Howaizi (Eaubonne Hospital); A. Queuniet (Elbeuf Hospital); P. Geoffroy, C. Picot (Epernay Private Hospital); J. Fournet, M. Mousseau (Grenoble University Hospital); C. Stampfli, P. M. Langlet (Laval Hospital); J. Doll (Le Chesnay University Hospital); A. Mignot; S. Durand (Le Havre Hospital); C. Buffet (Le Kremlin Bicêtre University Hospital); J. P. Triboulet, P. Quandalle, F. Denimal, M. Hebbar (Lille University Hospital); Mirabel (Lille Anticancer Centre Oscar Lambret); G. Lledo (Lyon Private Hospital St Jean); M. Giovannini (Marseille Anticancer Centre Paoli-Calmettes); P. Souillac (Meaux Hospital); M. Untereiner (Metz Private Hospital); C. Bernard (Metz Hospital); E. Leroy-Terquem (Meulan Hospital); H. Lacroix (Nantes University Hospital); E. Francois (Nice Anticancer Centre Antoine-Lacassagne); J. P. Lagasse, N. Breteau (Orléans Hospital); J. L. Legoux (Pessac University Hospital); J. C. Etienne (Poissy Hospital); J. F. Delattre, D. Lubrano, N. Levy-Chazal, J. P. Palot (Reims University Hospital); S. Nasca, L. Demange, T. D. Nguyen (Reims Anticancer Centre); J. Godinot, S. Seng (Rouen Anticancer Centre); H. Becquerel, P. Teniere (Rouen University Hospital); P. Thevenet, H. Le Brise, J. Fleury (Saint-Malo Hospital); J. Kammerer, H. Cosme (Sens Hospital); P. Novello (Saint Denis Anticancer Centre René-Huguenin); J. P. Avignon, P. Parisot, C. Berton, Legueul, G. Aunis (St Jean de Braye Private Hospital); D. Vetter (Strasbourg University Hospital); C. Platini (Thionville Hospital); L. Cals (Toulon Hospital); D. Rouhier, B. Robin, T. Champetier, A. Cartalat (Valence Private Hospital Pasteur/Générale/Polyclinique); C. Marchal, F. Guillemin (Vandoeuvre Anticancer Centre); A. Vautrin, M. Flamenbaum, D. Cassan (Vichy Hospital); M. Ducreux (Villejuif Anticancer Centre Gustave-Roussy).

references

1. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003; 349: 2241–2252.
2. Wu PC, Posner MC. The role of surgery in the management of oesophageal cancer. The Lancet Oncology 2003; 4: 481–488.

3. Blazeby JM, Farndon JR, Donavan J et al. A prospective longitudinal study examining the quality of life of patients with esophageal carcinoma. Cancer 2000; 88: 1781–1787.
4. Blazeby JM, Willams MH, Brookes ST et al. Quality of life measurement in patient with oesophageal cancer. Gut 1995; 37: 505–508.
5. Micallef J, Macquart-Moulin G, Auquier P et al. Quality of life assessment after surgery for oesophageal carcinoma : a critical review of the literature. Bull Cancer 1998; 85: 644–650.
6. Algan O, Coia LR, Keller SM et al. Management of adenocarcinoma of the esophagus with chemoradiation alone or chemoradiation followed by esophagectomy: results of sequential nonrandomized phase II studies. Int J Radiat Oncol Biol Phys 1995; 32: 753–761.
7. Denham JW, Burmeister BH, Lamb DS et al. Factors influencing outcome following radio-chemotherapy for oesophageal cancer. The Trans Tasman Radiation Oncology Group (TROG). Radiother Oncol 1996; 42: 91–92.
8. Chan A, Wong A. Is combined chemotherapy and radiation therapy equally effective as surgical resection in localized esophageal carcinoma? Int J Radiat Oncol Biol Phys 1999; 45: 265–270.
9. Hironaka S, Ohtsu A, Boku N et al. Non-randomized comparison between definitive chemoradiotherapy (CRT) and surgery (S) in patients (pts) with T2-3 M0 squamous cell carcinoma of the esophagus (SCE). ASCO Annual Meeting proceedings 2003; 20 : 159.
10. Siewert JR, Stein HJ, Feith M et al. Histologic tumor type is an independent prognostic parameter in esophageal cancer: lessons from more than 1,000 consecutive resections at a single center in the Western world. Ann Surg 2001; 234: 360–367.
11. Bedenne L, Michel P, Bouche O et al. Final results of a randomised multicentric phase III trial in locally advanced esophageal cancer: radiochemotherapy followed by surgery versus radiochemotherapy alone (FFCD 9102). Proc Am Soc Clin Oncol 2002; 21: 130a.
12. Sloan JA, Loprinzi CL, Kuross SA et al. Randomized comparison of four tools measuring overall quality of life in patients with advanced cancer. J Clin Oncol 1998; 16: 3662–3673.
13. Moinpour CM, Feigl P, Metch B et al. Quality of life end points in cancer clinical trials: review and recommendations. J Natl Cancer Inst 1989; 81: 485–495.
14. Lipscomb J, Donaldson MS, Arora NK et al. Cancer outcomes research. J Natl Cancer Inst Monogr 2004; 33: 178–197.
15. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. JAMA 1994; 272: 619–626.
16. Wood-Dauphinee S. Assessing quality of life in clinical research: from where have we come and where are we going? J Clin Epidemiol 1999; 52: 355–363.
17. Beitz J, Gnecco C, Justice R. Quality-of-life end points in cancer clinical trials: the U.S. Food and Drug Administration perspective. J Natl Cancer Inst Monogr 1996; 20: 7–9.
18. Johnson JR, Temple R. Food and Drug Administration requirements for approval of new anticancer drugs. Cancer Treat Rep 1985; 69: 1155–1159.
19. Homs MY, Essink-Bot ML, Borsboom GJ et al. Quality of life after palliative treatment for oesophageal carcinoma—a prospective comparison between stent placement and single dose brachytherapy. Eur J Cancer 2004; 40: 1862–1871.
20. de Boer AG, van Lanschot JJ, van Sandick JW et al. Quality of life after transhiatal compared with extended transthoracic resection for adenocarcinoma of the esophagus. J Clin Oncol 2004; 22:4202–4208.
21. Miller AB, Hoogstraten B, Staquet M, Winkler A. Cancer. Cancer 1981; 47: 207–214.
22. Spitzer WO, Dobson AJ, Hall J et al. Measuring the quality of life of cancer patients. A concise QL-Index for use by physician. J Chron Dis 1981; 34: 585–597.
23. Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of heath-related quality of life. Qual Life Res 1993; 2: 369–395.
24. Addington-Hall J, Kalra L. Measuring quality of life. Who should measure quality of life? BMJ 2001; 322: 1417–1420.
25. Moinpour CM, Lyons B, Schmidt SP et al. Substituting proxy ratings for patient ratings in cancer clinical trials: an analysis based on a Southwest Oncology Group trial in patients with brain metastases. Qual Life Res 2000; 9: 219–231.
26. Simes RJ, Greatorex V, Gebski VJ. Practical approaches to minimize problems with missing quality of life data. Stat Med 1998; 17: 725–737.
27. Bernhard J, Cella DF, Coates AS et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. Stat Med 1998; 17: 517–532.
28. Blazeby JM, Williams MH, Alderson D et al. Observer variation in assessment of quality of life in patients with oesophageal cancer. Br J Surg 1995; 82: 1200–1203.
29. Bernhard J, Sullivan M, Hurny C et al. Clinical relevance of single item quality of life indicators in cancer clinical trials. Br J Cancer 2001; 84: 1156–1165.
30. Sloan JA, Aaronson N, Cappelleri JC et al. Assessing the clinical significance of single items relative to summated scores. Mayo Clin Proc 2002; 77: 479–487.
31. Streiner DL, Norman GR. Health Measurement Scales. A practical guide to their development and use. Oxford: Oxford University Press 1989.
32. Dixon WJ, Merdian K. ANOVA and Regression with BMDP 5V. Los Angeles 1992.
33. Diggle PJ. An approach to the analysis of repeated measurements. Biometrics 1988; 44: 959–971.
34. Cnaan A, Laird NM, Slasor P. Tutorial in biostatistics. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. Stat Med 1997; 16: 2349–2380.
35. Pater J, Osoba D, Zee B et al. Effects of altering the time of administration and the time frame of quality of life assessments in clinical trials: an example using the EORTC QLQ-C30 in a large anti-emetic trial. Qual Life Res 1998; 7: 273–278.
36. Blazeby JM, Brookes ST, Alderson D. The prognostic value of quality of life scores during treatment for oesophageal cancer. Gut 2004; 49: 227–230.
37. Chau I, Norman AR, Cunningham D et al. Multivariate prognostic factor analysis in locally advanced and metastatic esophago-gastric cancer—pooled analysis from three multicenter, randomized, controlled trials using individual patient data. J Clin Oncol 2004; 22: 2395–2403.
38. Van Knippenberg FC, Out JJ, Tilanus HW et al. Quality of life in patients with resected oesophageal cancer. Soc Sci Med 1992; 35: 139–145.
39. Osoba D, Rodrigues G, Myles J et al. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998; 16: 139–144.
40. Stahl M, Stuschke M, Lehmann N et al. Chemoradiation with and without surgery in patients with locally advanced squamous cell carcinoma of the esophagus. J Clin Oncol 2005; 23: 2310–2317.
41. Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. Boca Raton: Chapman and Hall 2002.
42. Fairclough DL, Peterson HF, Cella D et al. Comparison of several model-based methods for analysing incomplete quality of life data in cancer clinical trials. Stat Med 2004; 17: 781–796.
43. Pauler DK, McCoy S, Moinpour C. Pattern mixture models for longitudinal quality of life studies in advanced stage disease. Stat Med 2003; 22: 795–809.
44. Zee BC. Growth curve model analysis for quality of life data. Stat Med 1998; 17: 757–766.
45. Bonnetain F, Bouche O, Conroy T et al. Longitudinal quality of life study in patients with metastatic gastric cancer. Analysis modalities and clinical applicability of QoL in randomized phase II trial in a digestive oncology? Gastroenterol Clin Biol 2005; 29: 1113–1124.
46. Sneeuw KC, Aaronson NK, Sprangers MA et al. Value of caregiver ratings in evaluating the quality of life of patients with cancer. J Clin Oncol 1997; 15: 1206–1217.
47. Sneeuw KC, Aaronson NK, Sprangers MA et al. Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. J Clin Epidemiol 1998; 51: 617–631.
48. Sneeuw KC, Aaronson NK, Sprangers MA et al. Evaluating the quality of life of cancer patients: assessments by patients, significant others, physicians and nurses. Br J Cancer 1999; 81: 87–94.
49. Blazeby JM, Alderson D, Winstone K et al. Development of an EORTC questionnaire module to be used in quality of life assessment for patients with oesophageal cancer. The EORTC Quality of Life Study Group. Eur J Cancer 1996; 32A: 1912–1917.
50. Blazeby JM, Conroy T, Hammerlid E et al. Clinical and psychometric validation of an EORTC questionnaire module, the EORTC QLQ-OES18, to assess quality of life in patients with oesophageal cancer. Eur J Cancer 2003; 39: 1984–1994.
51. Wiklund I, Karlberg J. Evaluation of quality of life in clinical trials. Selecting quality-of-life measures. Control Clin Trials 1991; 12: 204S–216S.