Quality of life of palliative chemotherapy naive patients with advanced adenocarcinoma of the stomach or esophagogastric junction treated with irinotecan combined with 5-fluorouracil and folinic acid: results of a randomised phase III trial

Desmond Curran · Carmelo Pozzo · Jerzy Zaluski · Magdalena Dank ·
Carlo Barone · Vahur Valvere · Suayib Yalcin · Christian Peschel ·
Miklós Wenczl · Erdem Goker · Roland Bugat

Abstract

Purpose The quality of life (QL) of advanced gastric cancer patients receiving irinotecan, folinic acid and 5-fluorouracil (5-FU) (IF arm) or cisplatin with 5-FU (CF arm) is presented.

Methods Patients with measurable or evaluable advanced gastric cancer received IF weekly for 6/7 weeks or CF q4 weeks. QL was assessed using the EORTC QLQ-C30 at baseline, subsequently every 8 weeks until progression and thereafter every 3 months until death. The QL data were analysed using several statistical methods including summary measures and pattern-mixture modelling.

Results A total of 333 patients were randomised and treated (IF 170, CF 163). The time-to-progression for IF and CF was 5.0 and 4.2 months ($P = 0.088$), respectively. The overall compliance rates for QL questionnaire completion were 60 and 56% in the IF and CF arms, respectively. Significant treatment differences were observed for the physical functioning scale ($P = 0.024$), nausea\vomiting ($P = 0.001$) and EQ-5D thermometer ($P = 0.020$) in favour of the IF treatment arm.

Conclusion There was a trend in favour of IF over CF in time-to-progression. The IF group also demonstrated a better safety profile than CF and a better QL on a number of multi-item scales, suggesting that IF offers an alternative first-line platinum-free treatment option for advanced gastric cancer.

Keywords Quality of life · Advanced gastric cancer · EORTC QLQ-C30 · EuroQoL · Irinotecan · CPT-11

Introduction

Gastric cancer is the second most common cause of cancer mortality worldwide [1]. Patients with advanced disease represent over two-thirds of newly diagnosed cases [2].

D. Curran (✉)
Omega Research, 189 Shanliss Rd., Santry, Dublin 9, Ireland
e-mail: currandes@omega-research.eu.com;
currandes@yahoo.com

C. Pozzo · C. Barone
Catholic University of Sacred Heart, 00100 Rome, Italy

J. Zaluski
Wielkopolskie Centrum Onkologii Poznan, Poznan, Poland

M. Dank
Semmelweis University, Budapest 78/a, Hungary

V. Valvere
Estonian Oncology Center, EE0016 Tallinn, Estonia

S. Yalcin
Medical Faculty, Institute of Oncology, Hacettepe University, 06100 Sihhiye, Ankara, Turkey

C. Peschel
Technical University Munich, 81675 Munich, Germany

M. Wenczl
Markusovszky County Hospital, Markusovszky u.3, Szombathely, Hungary

E. Goker
Ege University Medical School, Izmir, Turkey

R. Bugat
Institut Claudius Regaud, 31000 Toulouse, France

Despite advances in diagnosis and treatment, the prognosis for patients with advanced disease remains poor with the median survival reported to range from 5.3 to 10.2 months [3]. Several randomised trials demonstrated survival and quality of life (QL) benefits of chemotherapy compared with best supportive care [4]. Fluorouracil (5-FU) remains one of the most widely used drugs with the introduction of several other agents including cisplatin and anthracyclines being investigated in doublet and triplet combinations together with 5-FU or capcitabene [5–7]. The survival advantage of any of these combination regimens, compared with each other, is small and as such no internationally accepted standard of care regimen has emerged [6, 8]. The primary objectives of treatment in this palliative setting are to relieve symptoms, improve QL and prolong survival [9, 10]. However, a recent literature review and meta-analysis concluded that the impact of chemotherapy-related toxicity on the patients' quality of life has been insufficiently studied in patients with advanced gastric cancer [7].

Webb et al. compared the combination of epirubicin, cisplatin and 5-FU (ECF) with 5-FU, doxorubicin and methotrexate (FAMTX) in previously untreated patients with advanced esophagogastric cancer [11]. Using the EORTC QLQ-C30, the authors showed that ECF resulted in a better QL at 24 weeks compared with FAMTX. Subsequently, Ross et al. (2002) showed that ECF resulted in a better QL at 3 and 6 months when compared with mytomycin C, cisplatin and 5-FU (MCF) [12]. ECF has never been directly compared with CF, although a meta-analysis suggests ECF has a survival advantage over CF. However, concerns remain over the toxicity of ECF and the role of epirubicin in the combination [13]. More recently, Van Cutsem et al. (2006) compared the combination treatments docetaxel plus cisplatin and fluorouracil (DCF) vs. cisplatin and fluorouracil (CF) as first-line therapy for advanced gastric cancer [14]. The study met its primary endpoint showing a significant improvement in time-to-progression (TTP) with DCF compared with CF, and also an improvement in survival, and response rate were reported. Although, higher incidences of toxicity were observed in the DCF treatment arm, this did not appear to impact on QL which was significantly better in the DCF arm. These results suggest that better tumour control may also have lead to better symptom control in the DCF arm [14, 15].

Following promising results using irinotecan in combination with either 5-FU or cisplatin in phase II trials [16–19], a phase II–III trial was initiated in previously untreated advanced gastric adenocarcinoma patients comparing irinotecan plus cisplatin with irinotecan plus 5-FU given as an infusional AIO regimen (IF) [20]. Based on the risk/benefit ratio for IF in this study, a phase III trial was designed to compare IF to cisplatin plus 5-FU administered using a 5-day infusional regimen (CF). QL results for the phase III part of the study are reported here. The clinical results and initial QL results have been presented elsewhere [21].

As with other studies [14], the initial analysis of the QL data for the current study was carried out using time-to-event analysis (e.g. time to a 5% deterioration of the global health status/QL scale) in accordance with the statistical analysis plan. It is generally considered that for QL data, time-to-event analysis is limited since it does not take into account the repeated measures aspect of the data or the potential bias introduced by missing data. The analysis presented in this report addresses the fact that QL is a process and consequently is subject to change over time, that measurements taken at different time points are correlated, and that patients drop out during the study or have intermittent missing data, thus taking the entire structure of the QL data into account.

Patients and methods

Patient eligibility

Patients were to have histologically confirmed adenocarcinoma (including diffuse type, intestinal type and linitis) of the stomach or esophagogastric junction, with measurable or evaluable metastatic disease (cytology or histology was mandatory if a single metastatic lesion was the only manifestation of disease) or locally recurrent disease with at least one measurable lymph node. Patients were also required to be between 18 and 75 years of age, have Karnofsky performance status (KPS) >70%, life expectancy >3 months and adequate haematological parameters. The study was conducted in accordance with the Declaration of Helsinki and was approved by national or local ethics committees, as appropriate. All patients provided written informed consent. Further details regarding patient eligibility are provided elsewhere [21].

Study treatments

Subjects randomised to the IF arm were scheduled to receive irinotecan 80 mg/m^2 over a 30-min i.v. infusion, followed by FA 500 mg/m^2 as a 2-h i.v. infusion, immediately followed by 5-FU 2,000 mg/m^2 over a 22-h i.v. infusion, day 1 every week for 6 weeks followed by a 1-week rest. In the CF arm patients were scheduled to receive cisplatin 100 mg/m^2 as a 1–3-h i.v. infusion, day 1, followed by 5-FU 1,000 mg/m^2/day over a 24-h i.v. infusion, days 1–5, and every 4 weeks. Treatment was administrated until disease progression, unacceptable toxicity or withdrawal of consent.

Study design

The primary objective of the phase III part of the study was to detect a statistically significant increase in TTP for the IF arm relative to the CF arm. In addition, a non-inferiority comparison was specified in the protocol, in case of a non-significant trend towards superiority of TTP for the IF arm [22, 23]. Tumour progression was assessed according to World Health Organization Criteria and TTP measured from randomisation until the date of progression or death. Randomisation was performed using the minimisation technique [24], stratifying patients according to presence of measurable vs. evaluable disease, liver involvement, baseline weight loss, prior surgery and centre.

QL assessments

The EORTC QLQ-C30 (version 3.0) and EuroQoL (EQ-5D) instruments were used in this study. The QLQ-C30 is a cancer specific, self administered assessment of QL. The scale scores were calculated as per the scoring procedure defined in the EORTC QLQ-C30 Scoring Manual [25]. The EQ-5D is also a self-administered instrument comprising five questions and a visual analogue scale, which represents a rating of the patient's health state [26]. The five single items are combined to obtain a health utility index (HUI) score. This report focuses on the global health status\QL, physical functioning, social functioning, pain and nausea/vomiting scales of the EORTC QLQ-C30 and the two EQ-5D scales.

Quality of life assessments were required at baseline, every 8 weeks until documentation of disease progression and every 3 months from the documentation of the progression until death. To be considered evaluable at baseline, a questionnaire must have been filled in within 15 days before randomisation. To be considered evaluable on treatment, a questionnaire had to be filled in more than 5 days (IF arm) or 9 days (CF arm) after the start of the latest infusion so as not to take into account the immediate toxicities following infusion. The different lag durations after the start of the infusion allowed for the different infusion durations to be taken into account (1 day in the IF arm, 5 days in the CF arm).

Questionnaires without a date of assessment, or filled in after the cut-off date or after a further anti-tumour therapy were excluded (i.e. considered non-evaluable). Since assessments were planned to be evaluated independently from cycle duration, data were to be analysed according to time windows (8 week periods, i.e. ±4 weeks of the theoretical assessment date for assessments before investigator documented progressive disease). In case of multiple evaluable questionnaires in a time window, the mean value per subject for each scale in the time window was calculated.

Questionnaires excluded from the analysis were either considered present but not evaluable (i.e. see above description) or missing. The reason for missing questionnaires was collected on the CRF pages. The reasons were categorised as follows: random (i.e. administrative and similar reasons not directly related to patient QL), QL related (e.g. the patient was too ill to complete the questionnaire) or dead.

Statistical methods

Quality of life compliance was calculated as the ratio of the total number of subjects with at least one evaluable questionnaire per time window over the total number of expected questionnaires [27]. Patients were counted in the total number of expected questionnaires in the window only until further anti-tumour therapy or death prevented assessment.

Summary measures of QL scores were generated: i.e. the minimum, maximum and mean post-baseline QL scores within each patient, for each scale over all evaluable questionnaires, were calculated and summarised by treatment group. The Wilcoxon non-parametric test was used to compare treatment groups as the summary measures, particularly for the minimum and maximum, were not expected to be normally distributed.

A logistic regression model was fitted to test if the QL data missing from patients who dropped out was missing completely at random (MCAR) [28]. The model included terms for time (as a linear variable expressing the 8-weekly assessment time points), treatment (as a binary variable), time by treatment interaction and two terms representing global health status\QL scores: (1) sum of the two previous scores and (2) the difference between the two previous scores. The P-value for the Wald chi-squared statistic was used to test the effect of QL scores on dropout.

A pattern-mixture model was fitted in SAS using Proc Mixed [29, 30]. This model allows one to model the repeated measures structure of the data taking into account the dropout pattern. Terms in the model included treatment, time, dropout pattern and their interactions. Thus, a priori, the fixed effects as well as the covariance parameters were allowed to vary unconstrained according to the dropout pattern. In addition, several baseline clinical variables (age, gender, WHO performance status, pain assessed by the clinician, prior surgery and weight loss) were considered as covariates in the model. Model reduction was carried out using a likelihood ratio test to identify the most parsimonious model consistent with the data. If the treatment effect in the final model was pattern dependent, the delta method would be used to obtain the marginal treatment effect [31]. As such, the treatment effect is estimated within each pattern and the overall marginal treatment effect is

estimated using a weighted summation of the individual within pattern treatment effects, weighted according to the proportion of subjects in each dropout pattern. The null hypothesis of no treatment effect would be tested using a Wald statistic, which approximates a chi-squared distribution with one degree of freedom.

Results

Clinical results

Between June 2000 and March 2002, 337 patients were randomised (172 IF, 165 CF). Two patients in each arm were never treated, one due to disease progression and three due to adverse events. Thus the full analysis population, defined as treated patients analysed in the arm to which they were randomised, consisted of 333 patients (170 IF, 163 CF). The median treatment duration was 21 and 17 weeks in the IF and CF arms, respectively. The proportion of patients for whom an adverse event was reported as the reason for discontinuation was higher in the CF arm (10.0% IF, 21.5% CF; $P = 0.004$). A non-significant trend towards a longer median time-to-progression (TTP) was observed in the IF arm (5.0 months) compared with the CF arm (4.2 months: $P = 0.088$). The non-inferiority criteria, that the lower limit of the 95% CI of the Cox hazard ratio be at least 0.93, was satisfied for TTP in the full analysis population but not in the per protocol population. The median overall survival was 9.0 and 8.7 months in the IF and CF arms, respectively.

Safety was assessed in the 333 treated patients, according to treatment actually received (167 patients IF, 166 patients CF). A total of six treatment-related deaths occurred, one in the IF arm and five in the CF arm. The rate of hospitalisations was similar between arms, including the rate of hospitalisations due to toxicity (27.6% of patients). Neutropenia grade 3–4 was observed in 24.8 and 51.6% of IF and CF patients, respectively. Thrombocytopenia grade

3–4 was observed in 1.8 and 11.7% of IF and CF patients, respectively. Diarrhoea was observed more frequently in the IF arm (21.6 vs. 7.2% grade 3–4) whereas stomatitis was more prevalent in the CF arm (16.9 vs. 2.4%). Neurological toxicity was also more frequent in the CF arm, with 22.9% of patients experiencing grade 1–4 events, compared with 5.4% in the IF arm. Further clinical results are provided elsewhere [21].

QL compliance

Table 1 presents the compliance of the EORTC QLQ-C30 questionnaires by treatment group. The number of patients in each time window decreased over time due to attrition of patients. The compliance rate was higher in the IF arm at weeks 8, 16, 24, 40 and 48. The overall compliance rates were 60 and 56% in the IF and CF arms, respectively. Sixteen patients did not complete any evaluable questionnaires during this period. Monotone dropout patterns (i.e. a complete series of questionnaires before dropout) were observed in 202 cases. The remaining 115 patients had intermittent missing questionnaires. Of the 202 monotone dropout cases, 98 patients completed the baseline questionnaire only.

Table 2 presents the reasons for missing\non-evaluable questionnaires. During the first 48 weeks, a total of 63 and 57 assessments which were present were excluded from the analysis because of non-evaluability in the IF and CF treatment arms, respectively. The main reason for missing assessments at earlier visits (i.e. baseline, weeks 8 and 16) was due to administrative and similar reasons not directly related to patient QL, whereas the main reason for missing questionnaires at later time points was due to death.

Summary measures

The minimum, maximum and mean post-baseline QL scale scores were calculated for each patient and summarised by treatment group (see Table 3). There was no significant

Table 1 Analysis of compliance for QLQ-C30 questionnaires by protocol-planned assessment (full analysis population-randomisation group)

Time window (TW)	IF ($N = 170$)			CF ($N = 163$)		
	No. Pts in TW (N1)	No. Pts with at least one questionnaire (N2)	Rate: N2/N1 (%)	No. Pts in TW (N1)	No. Pts with at least one questionnaire (N2)	Rate: N2/N1 (%)
Baseline	170	145	85.3	163	143	87.7
Week 8	162	97	59.9	149	79	53.0
Week 16	138	76	55.1	126	57	45.2
Week 24	106	55	51.9	91	35	38.5
Week 32	73	27	37.0	63	26	41.3
Week 40	44	19	43.2	29	11	37.9
Week 48	29	14	48.3	16	3	18.8

*Time windows up to 1 year

Table 2 Reasons for missing\non evaluable questionnaires

Time window (TW)	Treatment	Missing\non-evaluable questionnaire	Reason				
			Present but non-evaluable	Random	QoL related	Dead	Unknown
Baseline	IF ($N = 170$)	25 (14.7)	13 (7.7)	7 (4.1)	0 (0.0)	0 (0.0)	5 (2.9)
	CF ($N = 163$)	20 (12.3)	7 (4.3)	12 (7.4)	0 (0.0)	0 (0.0)	1 (0.6)
Week 8	IF ($N = 170$)	73 (42.9)	5 (2.9)	39 (22.9)	13 (7.7)	8 (4.7)	8 (4.7)
	CF ($N = 163$)	84 (51.5)	6 (3.7)	52 (31.9)	8 (4.9)	7 (4.3)	11 (6.8)
Week 16	IF ($N = 170$)	94 (55.3)	13 (7.7)	44 (25.9)	9 (5.3)	17 (10.0)	11 (6.5)
	CF ($N = 163$)	106 (65.0)	9 (5.5)	49 (30.1)	13 (8.0)	23 (14.1)	12 (7.4)
Week 24	IF ($N = 170$)	115 (67.7)	8 (4.7)	39 (22.9)	7 (4.1)	41 (24.1)	20 (11.8)
	CF ($N = 163$)	128 (78.5)	13 (8.0)	37 (22.7)	8 (4.9)	52 (31.9)	18 (11.0)
Week 32	IF ($N = 170$)	143 (84.1)	13 (7.7)	30 (17.7)	12 (7.1)	73 (42.9)	15 (8.8)
	CF ($N = 163$)	137 (84.1)	10 (6.1)	29 (17.8)	9 (5.5)	75 (46.0)	14 (8.6)
Week 40	IF ($N = 170$)	151 (88.8)	8 (4.7)	28 (16.5)	4 (2.4)	98 (57.7)	13 (7.7)
	CF ($N = 163$)	152 (93.3)	9 (5.5)	31 (19.0)	7 (4.3)	99 (60.7)	6 (3.7)
Week 48	IF ($N = 170$)	156 (91.8)	3 (1.8)	26 (15.3)	2 (1.2)	116 (68.2)	9 (5.3)
	CF ($N = 163$)	160 (98.2)	3 (1.8)	40 (24.5)	3 (1.8)	110 (67.5)	4 (2.5)

difference in QL scores between treatment groups for the minimum global health status\QL scale. However, there was a non-significant trend towards a difference when comparing the maximum ($P = 0.062$) and mean global health status\QL scores ($P = 0.061$) between groups suggesting a trend towards a better QL for patients receiving IF. The physical functioning scale and the EQ-5D thermometer consistently presented significantly better results for all summary measures in favour of the IF treatment arm. The nausea\vomiting scale and EQ-5D HUI also indicated significant results for both the minimum and mean summary measures in favour of the IF treatment arm. Trends in favour of IF were also exhibited for the social functioning and pain scales.

Testing the dropout process

Table 4 presents the logistic regression analysis of dropout. The two QL terms in the model "difference in QL" and "sum of QL" were significant indicating that if the sum of the two previous QL scores were low then the probability of dropout was high and if there was a decrease in QL score from the previous assessment then the probability of dropout was also high. Thus, the missing data are not MCAR and caution needs to be taken when analysing the QL data.

Figure 1 presents the mean global health status\QL scores by time and dropout pattern for each treatment group. Dropout patterns were defined based on the time of the last completed questionnaire. Four patterns were defined as follows: 1 = dropout at baseline, 2 = dropout at week 8 or 16, 3 = dropout at week 24 or 32, 4 = dropout

after week 32. This resulted in 80, 88, 91 and 58 patients in the four patterns, respectively, with sufficient data within each pattern to carry out formal statistical analyses. Figure 1 illustrates that the mean global health status\QL score increased initially in all patterns in the IF treatment arm, however the mean scores subsequently decreased prior to dropout indicating that there was an initial improvement in global health status\QL score, possibly due to treatment, and a subsequent deterioration prior to dropout. In both treatment groups, there is a clear indication of differences between patterns with respect to mean global health status\QL score. The findings from Fig. 1 are consistent with the logistic regression analysis, in particular it illustrated that patients with a low QL score and a decrease in QL had a higher probability of dropout.

Model fitting

Several baseline clinical variables were considered as covariates in the model. The final model included an autoregressive order 1 variance–covariance structure, the baseline variables pain and WHO performance status and the treatment by pattern interaction and the main time effect. As there was an interaction between the treatment effect and pattern the treatment effect was estimated using the delta method. Figure 2 presents the treatment estimates for all the QL variables investigated except for the EQ5D HUI score ($P = 0.518$) which is on a different scale. Significant treatment differences were observed for the physical functioning scale, nausea\vomiting and EQ-5D thermometer in favour of the IF treatment arm. All the other scales illustrated non-significant results.

Table 3 Summary measures for secondary QL endpoints

Scale	Statistic	Summary measure	IF	CF	P-value
[a]Global health status\QL	N		116	101	
	Mean (SD)	Minimum	55.06 (21.90)	51.24 (21.19)	0.259
		Maximum	68.75 (21.92)	62.38 (24.05)	0.062
		Mean	62.41 (20.05)	56.95 (21.10)	0.061
[a]Physical functioning	N		117	101	
	Mean (SD)	Minimum	72.19 (23.47)	66.01 (24.57)	0.050
		Maximum	86.15 (15.98)	75.70 (23.14)	<0.001
		Mean	79.60 (17.68)	71.05 (22.55)	0.003
[a]Social functioning	N		116	102	
	Mean (SD)	Minimum	66.45 (29.52)	62.77 (28.90)	0.329
		Maximum	84.34 (20.88)	76.80 (28.04)	0.053
		Mean	76.28 (22.25)	70.62 (26.72)	0.164
[b]Pain	N		117	102	
	Mean (SD)	Minimum	13.96 (23.16)	19.44 (27.03)	0.092
		Maximum	31.48 (29.40)	31.73 (30.65)	0.970
		Mean	21.54 (23.24)	24.65 (26.51)	0.500
[b]Nausea\vomiting	N		116	102	
	Mean (SD)	Minimum	7.76 (16.07)	15.06 (22.32)	0.001
		Maximum	21.12 (22.56)	27.12 (27.40)	0.150
		Mean	13.62 (16.80)	20.82 (23.06)	0.026
[a]EQ-5D thermometer	N		87	69	
	Mean (SD)	Minimum	67.98 (19.18)	58.32 (18.68)	0.003
		Maximum	78.91 (16.10)	70.76 (19.08)	0.002
		Mean	73.66 (16.56)	64.80 (17.49)	0.002
[a]EQ-5D HUI	N		86	66	
	Mean (SD)	Minimum	0.69 (0.29)	0.58 (0.29)	0.003
		Maximum	0.82 (0.21)	0.73 (0.29)	0.110
		Mean	0.76 (0.23)	0.66 (0.27)	0.018

[a] A higher score represents a higher level of functioning and a better QL

[b] A higher score represents a higher level of symptom

Table 4 Logistic regression of dropout

Parameter	Estimate	Standard error	Chi-squared	P-value
Intercept	0.0146	1.0063	0.0002	0.9884
Time	−0.2350	0.2544	0.8535	0.3556
Treatment	−0.0857	0.3681	0.0542	0.8159
Time × treatment	0.1599	0.1097	2.1245	0.1450
Difference in QL	−0.0182	0.0064	8.2306	0.0041
Sum of QL	−0.0079	0.0029	7.1555	0.0075

Discussion

In this study, preliminary analysis using summary measures was carried out in an exploratory fashion. There were a number of significant results in the comparison of the two treatment groups consistently indicating a better QL in the IF treatment group. The main differences between treatment groups were observed for the physical functioning and nausea\vomiting scales from the EORTC QLQ-C30 and the two EQ-5D scales. Non-significant trends towards a difference were observed for the social functioning, pain and global health status\QL scale in favour of IF.

More questionnaires were completed in the IF arm than in the CF arm. As such the probability of observing an extreme result (e.g. minimum or maximum) is increased in the IF arm since the more frequently a process is observed the more often one will observe an extreme result. The number of questionnaires and the patterns of completion of questionnaires also varied considerably between patients. Missing data were prevalent. It was shown that missing data at earlier time points were due mainly to random reasons, e.g. administrative failure whereas missing questionnaires at later time points were missing mainly due to death. As
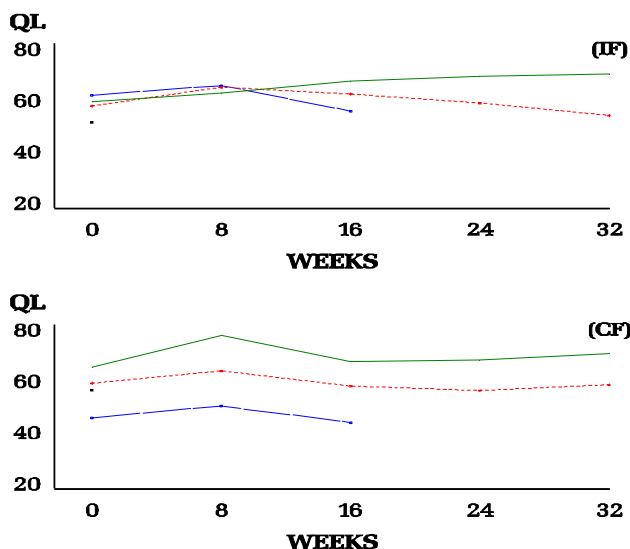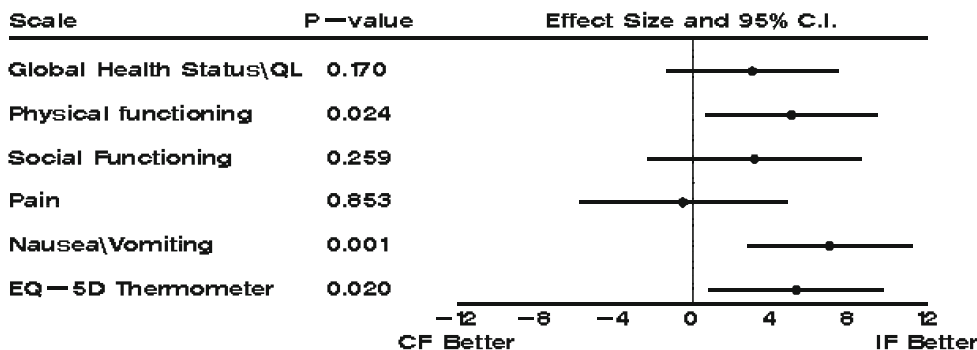
Fig. 1 *Plot* of the least squares means estimates of the EORTC QLQ-C30 mean global health status by treatment group and dropout pattern. A higher score represents a better QL. *Full green line* represents pattern 4 (i.e. dropout after week 32). *Other lines* represent patterns 1–3, i.e. dropout at baseline, dropout at week 8 or 16, dropout at week 24 or 32, respectively

Fig. 2 Testing the treatment effect using the delta method. EQ-5D HUI is on a different scale and consequently is not included in this figure

component. Analysing the data neglecting to take this information into account could be considered wasteful and potentially biased. Using the pattern-mixture model, significant treatment differences were observed for the physical functioning scale, nausea\vomiting and EQ-5D thermometer in favour of the IF treatment arm. These results were mainly consistent with those using the mean as a summary measure. However, for most scales the treatment effect was less significant using the pattern-mixture model. This is partially explained by the fact that between patient variation is artificially reduced using summary measures thus resulting in larger effect sizes. The findings of the QL analysis are also consistent with the toxicity profile as recorded through adverse event reporting [21]. While the rates of diarrhoea, cholinergic syndrome and fever without infection were higher in patients receiving IF, these symptoms were manageable in the current study. The higher rates of neurological toxicities, anorexia, stomatitis, alopecia, febrile neutropenia/neutropenic infection, thrombocytopenia and creatinine elevation in the CF arm, in addition to nausea and vomiting, are consistent with a negative impact on the physical functioning and nausea\vomiting scales.

such it may be argued that intermittent missing questionnaires were primarily due to random reasons (i.e. MCAR), whereas monotone missingness were due to progression of disease or death. This latter point is supported by testing the dropout process. Testing the dropout process indicated that questionnaires at the time of dropout were not MCAR. The results indicated that if QL scores were low then the probability of dropout was high. This phenomenon was confirmed when plotting the mean global health status\QL scores over time by dropout pattern. The imbalance in compliance and the dropout of patients suggests that simplified analyses such as time-to-event analysis and analysis using summary measures may be biased. Consequently, more complex analyses were carried out using pattern-mixture models to reduce any potential bias.

The final pattern-mixture model indicated that mean QL scores were dependent on dropout pattern and that the variance–covariance structure had an autoregressive

This was also reflected in the significantly higher number of withdrawals due to treatment-related AEs in the CF arm. In addition, the previously reported advantages in terms of efficacy (TTP and time to treatment failure) and clinical benefit (KPS, appetite and weight loss) all favoured patients receiving IF [21].

Analysis of the QL data using pattern-mixture analysis yielded more significant results than using time-to-event analysis. The original analysis of this study and other studies have used time-to-event analyses [14, 21]. However analyses of QL data using time-to-event are limited and potentially biased for a number of reasons. For example, when analysing the global health status\QL scale, 58% of patients had censored time-to-events. In analysis of QL data of advanced gastric patients where the time-to-event is time-to-deterioration of QL one could argue that there is informative censoring, i.e. missing questionnaires after dropout are not MCAR and consequently the probability of

being censored is not completely at random. This is particularly important in this study when analysing the global health status\QL scale as the majority of patients had censored time-to-events. Conversely, only 42% of patients had observed the event of interest (i.e. deterioration of QL score). As the number of QL events is small the power to detect a treatment difference is small. Consequently, even if large differences are expected between treatment groups, the probability of observing a significant difference is small. Thus time-to-event analyses would appear to be potentially biased and wasteful for analysing QL data.

Currently, there is no internationally agreed upon gold standard for conducting and reporting QL studies in cancer clinical trials [32]. While other authors have also used the EORTC QLQ-C30 questionnaire in advanced gastric cancer, sometimes reporting of results was poor and was limited to a few paragraphs within the overall clinical paper [11, 12, 14]. For example, details concerning compliance within treatment arms were not provided and methods of analysis were sub-optimal as they did not take into account the structure of the data, i.e. repeated (correlated) measurements with missing data. It is imperative that sufficient details concerning QL assessment, analysis and reporting are provided to allow comparisons of findings across studies. This is particularly relevant in diseases such as advanced gastric cancer where survival rates are similar across treatment regimens.

The use of irinotecan-based regimens for the treatment of advanced gastric cancer has been further explored in phase II studies during the last few years, especially with the availability of new targeted agents [33–35]. Although initial results are promising, suggesting that IF could represent a potential platinum-free alternative backbone to be combined with new targeted agents, results from phase III studies are required before drawing any firm conclusions. QoL assessment should be incorporated as a prominent objective in phase III studies in advanced gastric cancer to help both patients and physicians to discuss treatment choices and aid decision making [6, 7].

In summary, there was a trend in favour of IF over CF in time-to-progression. The IF treatment arm also demonstrated a better safety profile than the CF arm and a better QL on a number of multi-item scales. These results would suggest that IF offers an alternative platinum-free first-line treatment option for advanced gastric cancer which should be explored further in combination with new targeted agents.

References

1. Pisani, P., Parkin, D. M., Bray, F., & Ferlay, J. (1999). Estimates of the worldwide mortality from 25 cancers in 1990. *International Journal of Cancer, 83*(1), 18–29. doi:10.1002/(SICI)1097-0215(19990924)83:1<18::AID-IJC5>3.0.CO;2-M.

2. Jemal, A., Murray, T., Ward, E., Samuels, A., Tiwari, R. C., Ghafoor, A., et al. (2005). Cancer statistics, 2005. *CA: A Cancer Journal for Clinicians, 55*(1), 10–30. doi:10.3322/canjclin.55.1.10.

3. Dickson, J. L., & Cunningham, D. (2004). Systemic treatment of gastric cancer. *European Journal of Gastroenterology and Hepatology, 16*(3), 255–263. doi:10.1097/00042737-200403000-00003.

4. Bugat, R. (2003). Irinotecan in the treatment of gastric cancer. *Annals of Oncology, 14*(Suppl 2), S37–S40.

5. Van Cutsem, E. (2004). The treatment of advanced gastric cancer: New findings on the activity of the taxanes. *The Oncologist, 9*(Suppl 2), S9–S15.

6. Pozzo, C., & Barone, C. (2008). Is there an optimal chemotherapy regimen for the treatment of advanced gastric cancer that will provide a platform for the introduction of new biological agents? *The Oncologist, 13*(7), 794–806. doi:10.1634/theoncologist.2008-0082.

7. Wagner, A., Grothe, W., Haerting, J., Kleber, G., Grothey, A., & Fleig, W. E. (2006). Chemotherapy in advanced gastric cancer: A systematic review and meta-analysis based on aggregate data. *Journal of Clinical Oncology, 24*(18), 2903–2909. doi:10.1200/JCO.2005.05.0245.

8. Bouche, O., Raoul, J. L., Bonnetain, F., Giovannini, M., Etienne, P. L., Lledo, G., et al. (2004). Randomized multicenter phase II trial of a biweekly regimen of fluorouracil and leucovorin (LV5FU2), LV5FU2 plus cisplatin, or LV5FU2 plus irinotecan in patients with previously untreated metastatic gastric cancer: A Federation Francophone de Cancerologie Digestive Group Study—FFCD 9803. *Journal of Clinical Oncology, 22*(21), 4319–4328. doi:10.1200/JCO.2004.01.140.

9. Chau, I., Norman, A. R., Cunningham, D., Waters, J. S., Oates, J., & Ross, P. J. (2004). Multivariate prognostic factor analysis in locally advanced and metastatic esophago-gastric cancer-pooled analysis from three multicenter, randomized, controlled trials using individual patient data. *Journal of Clinical Oncology, 22*(12), 2395–2403. doi:10.1200/JCO.2004.08.154.

10. (1996). Outcomes of cancer treatment for technology assessment and cancer treatment guidelines: American Society of Clinical Oncology. *Journal of Clinical Oncology, 14*(2), 671–679.

11. Webb, A., Cunningham, D., Scarffe, J. H., Harper, P., Norman, A., Joffe, J. K., et al. (1997). Randomized trial comparing epirubicin, cisplatin, and fluorouracil versus fluorouracil,

doxorubicin, and methotrexate in advanced esophagogastric cancer. *Journal of Clinical Oncology, 15*(1), 261–267.

12. Ross, P., Nicolson, M., Cunningham, D., Valle, J., Seymour, M., Harper, P., et al. (2002). Prospective randomized trial comparing mitomycin, cisplatin, and protracted venous-infusion fluorouracil (PVI 5-FU) with epirubicin, cisplatin, and PVI 5-FU in advanced esophagogastric cancer. *Journal of Clinical Oncology, 20*(8), 1996–2004. doi:10.1200/JCO.2002.08.105.

13. Ilson, D. H. (2002). Epirubicin, cisplatin, and fluorouracil in gastric and esophageal cancer: A step ahead? *Journal of Clinical Oncology, 20*(8), 1962–1964.

14. Van Cutsem, E., Moiseyenko, V. M., Tjulandin, S., Majlis, A., Constenla, M., Boni, C., et al. (2006). Phase III study of docetaxel and cisplatin plus fluorouracil compared with cisplatin and fluorouracil as first-line therapy for advanced gastric cancer: A report of the V325 study group. *Journal of Clinical Oncology, 24*(31), 4991–4997. doi:10.1200/JCO.2006.06.8429.

15. Ajani, J. A., Moiseyenko, V. M., Tjulandin, S., Majlis, A., Constenla, M., Boni, X., et al. (2007). Quality of life with docetaxel plus cisplatin and fluorouracil compared with cisplatin and fluorouracil from a phase III trial for advanced gastric or gastroesophageal adenocarcinoma: The V-325 Study Group. *Journal of Clinical Oncology, 25*(22), 3210–3216. doi:10.1200/JCO.2006.08.3956.

16. Blanke, C., Haller, D., Benson, A. B., Rothenberg, M. L., Berlin, J., Mori, M., et al. (2001). A phase II study of irinotecan with 5-fluorouracil and leucovorin in patients with previously untreated gastric adenocarcinoma. *Annals of Oncology, 12*, 1575–1580. doi:10.1023/A:1013129315036.

17. Boku, N., Ohtsu, A., Shimada, Y., Shirao, K., Seki, S., Saito, H., et al. (1999). Phase II study of a combination of irinotecan and cisplatin against metastatic gastric cancer. *Journal of Clinical Oncology, 17*(1), 319–323.

18. Ajani, J. A. (2000). Standard chemotherapy for gastric carcinoma: Is it a myth? *Journal of Clinical Oncology, 18*(23), 4001–4003.

19. Moehler, M., Eimermacher, A., Siebler, J., Höhler, T., Wein, A., Menges, M., et al. (2005). Randomised phase II evaluation of irinotecan plus high-dose 5-fluorouracil and leucovorin (ILF) vs 5-fluorouracil, leucovorin, and etoposide (ELF) in untreated metastatic gastric cancer. *British Journal of Cancer, 92*(12), 2122–2128. doi:10.1038/sj.bjc.6602649.

20. Pozzo, C., Barone, C., Szanto, J., Padi, E., Peschel, C., Bükki, J., et al. (2004). Irinotecan in combination with 5-fluorouracil and folinic acid or with cisplatin in patients with advanced gastric or esophageal-gastric junction adenocarcinoma: Results of a randomized phase II study. *Annals of Oncology, 15*(12), 1773–1781. doi:10.1093/annonc/mdh473.

21. Dank, M., Zaluski, J., Barone, C., Valvere, V., Yalcin, S., Peschel, C., et al. (2008). Randomized phase III study of irinotecan combined with 5-fluorouracil and folinic acid and cisplatin combined with 5-fluorouracil in palliative chemotherapy naive patients with advanced adenocarcinoma of the stomach or esophagogastric junction. *Annals of Oncology, 19*(8), 1450–1457. doi:10.1093/annonc/mdn166.

22. Morikawa, T., & Yoshida, M. (1995). A useful testing strategy in phase III trials: Combined test of superiority and test of equivalence. *Journal of Biopharmaceutical Statistics, 5*(3), 297–306. doi:10.1080/10543409508835115.

23. Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

24. Freedman, J., & White, S. J. (1976). On the use of Pocock and Simon's method for balancing treatment numbers over prognostic factors in the controlled clinical trial. *Biometrics, 32*, 691–694. doi:10.2307/2529759.

25. Fayers, P. M., Aaronson, N. K., Bjordal, K., Curran, D., & Groenvold, M. (1999). *EORTC QLQ-C30 scoring manual* (2nd ed.). Brussels: EORTC.

26. Kind, P. (1996). The EuroQL instrument: An index of health-related quality of life. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 191–201). Philadelphia: Lippincott-Raven Publishers.

27. Machin, D., & Weeden, S. (1998). Suggestions for the presentation of quality of life data from clinical trials. *Statistics in Medicine, 17*(5–7), 711–724. doi:10.1002/(SICI)1097-0258(19980315/15)17:5/7<711::AID-SIM816>3.0.CO;2-H.

28. Curran, D., Molenberghs, G., Fayers, P. M., & Machin, D. (1998). Incomplete quality of life data in randomized trials: Missing forms. *Statistics in Medicine, 17*(5–7), 697–709. doi:10.1002/(SICI)1097-0258(19980315/15)17:5/7<697::AID-SIM815>3.0.CO;2-Y.

29. Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *Analysis of longitudinal data*. Oxford: Clarendon Press.

30. Verbeke, G., & Molenberghs, G. (1997). *Linear mixed models in practice*. New York: Springer–Verlag.

31. Curran, D., Molenberghs, G., Aaronson, N. K., Fossa, S. D., & Sylvester, R. J. (2002). Analysing longitudinal continuous quality of life data with dropout. *Statistical Methods in Medical Research, 11*(1), 5–23. doi:10.1191/0962280202sm270ra.

32. Efficace, F., & Bottomley, A. (2002). Do quality-of-life randomized clinical trials support clinicians in their decision-making? *Journal of Clinical Oncology, 20*(19), 4126. doi:10.1200/JCO.2002.99.117.

33. Moehler, M., Trarbach, T., Seufferlein, T., Kubicka, S., Lordick, F., Geissler, M., et al. (2007). Cetuximab with irinotecan/Na-FA/5-FU as first-line treatment in advanced gastric cancer. *European Journal of Cancer, 5*(4), Abstract 3518.

34. Pinto, C., Di Fabio, F., Siena, S., Cascinu, S., Rojas Llimpe, F. L., et al. (2007). Phase II study of cetuximab in combination with FOLFIRI in patients with untreated advanced gastric or gastroesophageal junction adenocancinoma (FOLCETUX study). *Annals of Oncology, 18*(3), 510–517. doi:10.1093/annonc/mdl459.

35. Shah, M. A., Ramanathan, R. K., Ilson, D. H., Levnor, A., D'Adamo, D., O'Reilly, E., et al. (2006). Multicenter phase II study of irinotecan, cisplatin, and bevacizumab in patients with metastatic gastric or gastroesophageal junction adenocarcinoma. *Journal of Clinical Oncology, 24*(33), 5201–5206.