Mean overall survival gain with aflibercept plus FOLFIRI *vs* placebo plus FOLFIRI in patients with previously treated metastatic colorectal cancer

F Joulain*,[1], I Proskorovsky[2], C Allegra[3], J Tabernero[4], M Hoyle[5], S U Iqbal[6] and E Van Cutsem[7]

[1]Sanofi, 1 av. Pierre Brossolette, Chilly Mazarin 91385, France; [2]Evidera, Dorval Ave, Suite 500, Dorval, Quebec, Canada H9S 5J9; [3]University of Florida/Shands, 1600 SW Archer Road, Box 100278, Gainesville, FL 32610, USA; [4]Vall d'Hebron University Hospital—Institute of Oncology (VHIO), P. Vall d'Hebron 119-129, Barcelona 08035, Spain; [5]University of Exeter, Peninsula Technology Assessment Group (PenTAG), Peninsula College of Medicine and Dentistry, Veysey Building, Salmon Pool Lane, Exeter EX2 4SG, UK; [6]Sanofi, 500 Kendall St, Cambridge, MA 02142, USA and [7]University Hospitals Gasthuisberg/Leuven, UZ Leuven, Campus Gasthuisberg, Herestraat 49, Leuven 3000, Belgium

Background: Mean survival in cancer trials can be estimated with statistical techniques to extrapolate study survival curves. This methodology was applied to data from the VELOUR trial, where use of the novel biologic aflibercept (ziv-aflibercept in the United States) in combination with fluorouracil + leucovorin + irinotecan (FOLFIRI), had significantly increased median overall survival (OS) by 1.44 months, *vs* placebo plus FOLFIRI in patients with metastatic colorectal cancer (mCRC) resistant to, or that had progressed following, an oxaliplatin-containing regimen.

Methods: Parametric survival analyses were used to identify distributions with the best fit to the empirical VELOUR data. Mean OS for the two treatment groups (and pre-defined subgroups) was calculated from the fitted curves over a 15-year survival period.

Results: Overall, the log-logistic distribution was the best-fitting for both treatment arms and, with it, the estimated difference in mean OS over 15 years between aflibercept + FOLFIRI and placebo + FOLFIRI was 4.7 months. In addition, the survival advantage with aflibercept was at least 3 months for the ITT population, whichever distribution was used to extrapolate survival.

Conclusion: Extrapolation of survival curves suggests the mean OS difference for aflibercept in the VELOUR trial is at least 3 months in the ITT population and selected subgroups.

Over 1.2 million people are diagnosed with colorectal cancer worldwide each year (Jemal *et al*, 2011; International Agency for Research on Cancer, 2011). Around a quarter of them will already have metastatic disease (metastatic colorectal cancer (mCRC)) and, consequently, over a 90% chance of dying within the next 5 years (Goldberg *et al*, 2007). Viewed in isolation, such data are daunting. However, a broader perspective must also embrace major therapeutic advances and consequent increases

in survival in non-resectable mCRC over the last 20 years (Jemal *et al*, 2011).

Currently, median overall survival (OS) in mCRC typically exceeds 2 years (Davies and Goldberg, 2008) using first-line treatments based on fluorouracil with leucovorin plus either oxaliplatin (FOLFOX) or irinotecan (FOLFIRI), often with additional biologic-targeted therapy. With FOLFOX and FOLFIRI in first- and second-line treatments, median survival of 20–21

months and mean survival of 24–27 months have been reported (Tournigand *et al*, 2004; Hind *et al*, 2008). Current real-world treatment patterns show that about half of patients progress and go on to receive second-line combination therapy (typically involving cross-over between FOLFIRI and FOLFOX) (Hess *et al*, 2010; Rosé *et al*, 2012), with similar patterns for use of additional biologic therapy in this setting. A new option for patients with mCRC resistant to or that has progressed following an oxaliplatin-containing regimen is aflibercept (ziv-aflibercept in the United States), a novel fusion protein that acts as a soluble receptor that binds to VEGF-A, VEGF-B and placental growth factor, added to FOLFIRI. In combination with FOLFIRI, aflibercept has been investigated in the VELOUR trial, a multinational, randomised, placebo-controlled Phase III study in patients who had progressed following an oxaliplatin-based regimen (Van Cutsem *et al*, 2012). Aflibercept demonstrated significant improvement in survival end points, extending both median OS (by 1.44 months) and median progression-free survival (PFS; by 2.23 months), at a median follow-up of 22.3 months in OS. These statistically significant survival benefits were observed together with an acceptable safety and tolerability profile.

Despite the established efficacy of aflibercept, the VELOUR results are limited by their reliance on median values of OS and PFS – standard measures in cancer intervention trials. On the basis of the shape of the OS curves, the median might not reflect the full benefit of aflibercept, as the curves for two treatment arms continue to separate, showing a higher benefit of aflibercept in the longer term.

As well as circumventing the problem that survival data in such studies are usually right-skewed, median survival outcomes (Davies *et al*, 2012) are also a pragmatic alternative to measuring mean survival directly. The latter approach (i.e., taking into account the whole of the survival curve, not just the survival of 50% of patients) would require waiting until all participants had died, which could raise ethical questions about whether it denied patients an effective and sufficiently safe treatment. It would also have major implications for research sponsors, given the reduced time for ensuring returns on investment for the investigated products before patent expiry. Against all this, however, is the fact that mean survival values (estimated separately or as an output of a cost-effectiveness analysis) are required by many payers and health technology assessment (HTA) agencies (e.g., the National Institute for Health and Care Excellence (NICE) in the United Kingdom or the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia) to estimate cost-effectiveness, to inform reimbursement decisions.

The more practical approach of relying on median OS overlooks people still alive long after study follow-up has finished, and hence risks not taking into account the appropriate clinical and associated health economic consequences of the interventions under investigation. A potential solution to this predicament is to use statistical techniques to extrapolate study survival curves to predict long-term mean outcomes beyond the observed follow-up time, an established method in health economic analyses (Latimer, 2011) and applied to data from numerous cancer trials (Neymark *et al*, 2002; Huse *et al*, 2007). The objective of this study was to assess the implications of such survival estimations in the context of the VELOUR trial by generating mean OS differences between the study's treatment groups using parametric survival analyses.

METHODS

The current study was based on *post hoc*-extended analysis of data from the VELOUR trial (full clinical results of which have been published elsewhere) (Van Cutsem *et al*, 2012), and was conducted to estimate the difference in mean survival between the trial's treatment groups.

Data. The VELOUR trial was a prospective multinational, double-blind, placebo-controlled study in which patients with mCRC who had had disease progression following previous treatment with an oxaliplatin-based regimen were randomised to receive either aflibercept + FOLFIRI or placebo + FOLFIRI (control). To be eligible for the study, patients had to be aged 18 years or older with histologically- or cytologically-proven colorectal adenocarcinoma with metastatic disease not amenable to potentially curative treatment, and must not have been treated with irinotecan prior to the trial entry. Participants were randomised (with stratification according to prior therapy with bevacizumab (yes *vs* no) and ECOG performance status (PS) (0 *vs* 1 *vs* 2)) to receive aflibercept (4 mg kg^{-1} intravenously) or placebo, every 2 weeks in combination with FOLFIRI, in a 1 : 1 ratio. They were treated until disease progression, unacceptable toxicity, patients' refusal, or investigators' decision to withdraw treatment. Following the documentation of progressive disease, patients were followed for survival status.

A total of 1226 patients were enrolled, 612 being randomised to aflibercept + FOLFIRI and 614 to placebo + FOLFIRI. Patient characteristics and disease history are summarised in Table 1. Median follow-up time for the intention-to-treat (ITT) population was 22.28 months, with 403 deaths being observed in the aflibercept arm and 460 in the control arm. Median reported OS was 12.06 months (95.34% CI: 11.07–13.11) with placebo + FOLFIRI and 13.50 months (95.34% CI: 12.52–14.95) with aflibercept + FOLFIRI (HR: 0.817; 95.34% CI: 0.713–0.937; $P = 0.0032$). Median reported PFS was also significantly longer with aflibercept + FOLFIRI than with placebo + FOLFIRI (6.90 *vs* 4.67 months, $P < 0.001$; HR: 0.758; 95% CI: 0.661–0.869). In addition, response rate was significantly higher in the aflibercept arm than in the control arm (19.8% *vs* 11.1%, $P = 0.0001$) (Van Cutsem *et al*, 2012).

The adverse events reported with aflibercept and FOLFIRI in the trial included the characteristic antivascular endothelial growth factor effects and also reflected an increased incidence of some chemotherapy-related toxicities (Van Cutsem *et al*, 2012). Treatment-emergent adverse events were reported in 99.2% and 97.9% of the aflibercept and control arm patients, respectively, with grade 3 and 4 events reported in 83.5% and 62.5% of patients, respectively. In addition, the reported incidence of some adverse events commonly associated with chemotherapy was higher with aflibercept, including the following grade 3 and 4 events: diarrhoea, asthenic conditions, stomatitis and ulceration, infections and palmar-plantar erythrodysesthesia (Van Cutsem *et al*, 2012). Adverse events led to permanent discontinuation from study treatment in 26.8% of patients in the aflibercept arm and 12.1% of patients in the control arm. The trends observed for adverse events in the subgroups were generally consistent with those of the ITT population (Van Cutsem *et al*, 2012).

Statistical analyses. At the end of 36 months of follow-up, 17.2% of patients in the aflibercept arm and 7.9% of patients in the placebo arm were still alive. Therefore, to derive mean OS, survival curves had to be extrapolated (Latimer, 2011). Parametric survival analyses were used to identify distributions that provided the best fit to the OS empirical data. First, observed Kaplan–Meier curves and cumulative hazard functions were examined graphically. If the exploratory analysis showed that the shape of the cumulative hazard was similar in both the treatment arms, then the two arms were modelled together and a treatment indicator was included as a predictor in the model; otherwise, each treatment arm was modelled separately. Commonly used distributions including exponential, Weibull, log-normal and log-logistic curves were tested, and the fit of each was assessed both by using statistical criteria (i.e., Akaike information criterion (AIC) and the Bayesian

Table 1. Patient characteristics and disease history

Characteristic	Placebo/FOLFIRI (N = 614)	Aflibercept/ FOLFIRI (N = 612)
ECOG PS, n (%)[a]		
0	350 (57.0)	349 (57.0)
1	250 (40.7)	250 (40.8)
2	14 (2.3)	13 (2.1)
Prior bevacizumab, n (%)[a]		
Yes	187 (30.5)	186 (30.4)
No	427 (69.5)	426 (69.6)
Gender, n (%)		
Male	353 (57.5)	365 (59.6)
Female	261 (42.5)	247 (40.4)
Age, years		
Median	61.0	61.0
Range	19–86	21–82
Primary site, n (%)		
Colon	302 (49.2)	289 (47.2)
Rectosigmoid	136 (22.1)	123 (20.1)
Rectum	174 (28.3)	197 (32.2)
Other	2 (0.3)	3 (0.5)
No. of metastatic organs involved at baseline (excluding primary site), n (%)		
0	6 (1.0)	2 (0.3)
1	271 (44.1)	256 (41.8)
>1	337 (54.9)	354 (57.8)
Metastatic organs involved at baseline (excluding primary site), n (%)		
Any site	608 (99.0)	610 (99.7)
Liver	431 (70.2)	459 (75.0)
Lung	277 (45.1)	271 (44.3)
Lymph nodes	181 (29.5)	173 (28.3)
Peritoneum	88 (14.3)	68 (11.1)
Liver metastasis, n (%)		
None, or liver and other metastases	468 (76.2)	459 (75.0)
Liver metastasis only	146 (23.8)	153 (25.0)
Prior hypertension, n (%)	268 (43.6)	266 (43.5)
Prior chemotherapy, n (%)		
Adjuvant only[b]	64 (10.4)	60 (9.8)
Adjuvant and metastatic disease	108 (17.6)	102 (16.7)
Metastatic disease	442 (72.0)	450 (73.5)

Abbreviations: ECOG PS = Eastern Cooperative Oncology Group performance status; FOLFIRI = irinotecan, leucovorin and infusional fluorouracil.
[a]ECOG PS and prior bevacizumab status as provided at randomisation (stratification variable).
[b]Relapse within 6 months of completion of oxaliplatin-based adjuvant therapy.

Table 2. Commonly used survival distributions in economic evaluation

	S(t)	Mean
Exponential	$\mathrm{Exp}\{-\lambda t\}$	$\dfrac{1}{\lambda}$
Weibull	$\mathrm{Exp}\{-\lambda t^{\gamma}\}$	$\left(\dfrac{1}{\lambda}\right)^{1/\gamma}\Gamma\left(\dfrac{1}{\gamma}+1\right)$
Lognormal	$1-\varnothing\left(\dfrac{\ln(t)-(\beta_0+X_J\beta_x)}{\sigma}\right)$	$\dfrac{1}{\psi}e^{\sigma^2/2}$
Loglogistic	$\left[1+\{\psi t\}^{1/\gamma}\right]^{-1}$	$\dfrac{1}{\psi}\dfrac{\gamma\pi}{\sin(\gamma\pi)}$

$\lambda = \exp(-(\beta_0 + X_J\beta_x)\gamma)$. $\Psi = \exp(-\beta_0 - X_J\beta_x)$. γ = shape parameter. Γ = gamma function. $\varnothing$ = cumulative standard normal distribution. Expression for log-logistic mean applies to $\gamma < 1$.

used as the first choice in modelling cancer survival (Muszbek *et al*, 2012).

To calculate mean OS from the fitted survival curve, two different approaches were used. The first used the closed-form solutions presented in Table 2 for calculating the area under the extrapolated survival curve for selected distributions. In the second approach, a survival cutoff time was applied by forcing the predicted survival curve to 0 at different time points (i.e., 5, 10 or 15 years) to limit the effect of the long tail of the log-logistic or log-normal distribution. The area under the truncated curve was then calculated using the Riemann sum approximation technique (i.e., dividing the area under a curve into small rectangles, calculating the area of each rectangle, and then summing up the areas to approximate the area under the curve (Shilov *et al*, 1977)). This second approach is commonly used in economic evaluations when information on maximum survival is lacking and the extrapolated survival curve has a long tail. Mean OS in patients treated with aflibercept + FOLFIRI *vs* patients treated with placebo + FOLFIRI alone were calculated and used to derive the difference in mean OS between the two treatment groups.

Analyses were performed for the entire ITT population of the VELOUR trial, as well as for the pre-defined subgroups of prognostic interest, including ECOG PS0, no prior bevacizumab, liver metastasis only and number of organ metastases ⩽1. Additional analyses were also done for ECOG PS 1 and prior bevacizumab subgroups.

RESULTS

Fit of statistical distributions. According to AIC and BIC, the log-logistic distribution was the best-fitting curve for the ITT population and all subgroups for the aflibercept arm, and for ITT and two subgroups for the placebo arm (Table 3). In the placebo arm, the log-normal distribution provided the best fit for three subgroups, followed by the log-logistic distribution; however, differences in predicted survival times between the log-normal and log-logistic models were minimal. In the placebo arm, for prior bevacizumab subgroup, the Weibull distribution provided the best fit.

Although the log-logistic distribution was found to provide the best fit for all patients combined for both the treatment arms, the shape of the curve (as determined by the shape parameter) differed between the aflibercept and control arms. This finding was supported by the observation that the ln(S/(1 − S)) *vs* ln(time) functions did not stay parallel and started to diverge towards the

information criterion (BIC) (Cleves *et al*, 2002; Singer and Willett, 2003) and graphically by comparing the empirical and predicted curves. Long-term predictions from the best-fitting distribution were examined for clinical plausibility and compared with the long-term prediction from the Weibull distribution. The Weibull distribution was chosen for such comparisons because it is often

Table 3. Statistical fit of common distributions

	ITT		ECOG PS 0		No prior bevacizumab		Liver Mets		One or less organ metastases		Prior bevacizumab use		ECOG PS 1	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Aflibercept														
Weibull	1497.1	1505.9	763.4	771.1	1056.9	1065.1	354.2	360.2	571.0	578.1	439.7	446.1	634.9	641.9
Log-normal	1491.1	1500.0	752.0	759.7	1056.6	1064.7	348.5	354.6	568.7	575.8	434.3	440.8	626.1	633.2
Log-logistic	**1479.7**	**1488.5**	**750.0**	**757.7**	**1048.0**	**1056.1**	**346.0**	**352.0**	**564.9**	**572.0**	**432.0**	**438.5**	**625.0**	**632.0**
Exponential	1535.4	1539.8	802.2	806.1	1077.3	1081.3	367.8	370.8	595.9	599.5	458.6	461.9	651.7	655.2
Placebo														
Weibull	1469.3	1478.2	755.0	762.7	997.0	1005.1	347.3	353.3	654.2	661.5	**473.3**	**479.8**	630.0	637.0
Log-normal	1469.1	1477.9	**735.5**	**743.2**	984.4	992.5	341.4	347.4	642.9	650.1	478.3	484.8	633.7	640.7
Log-logistic	**1457.2**	**1466.0**	739.2	746.9	**978.4**	**986.5**	343.8	349.8	647.9	655.2	475.8	482.3	**627.8**	**634.9**
Exponential	1549.7	1554.2	823.6	827.4	1065.6	1069.7	369.2	372.2	684.8	688.5	485.9	489.1	657.8	661.3

Abbreviations: AIC = Akaike information criterion; BIC = Bayesian information criterion; ITT = intention-to-treat. Bold values indicate best-fitting distribution.
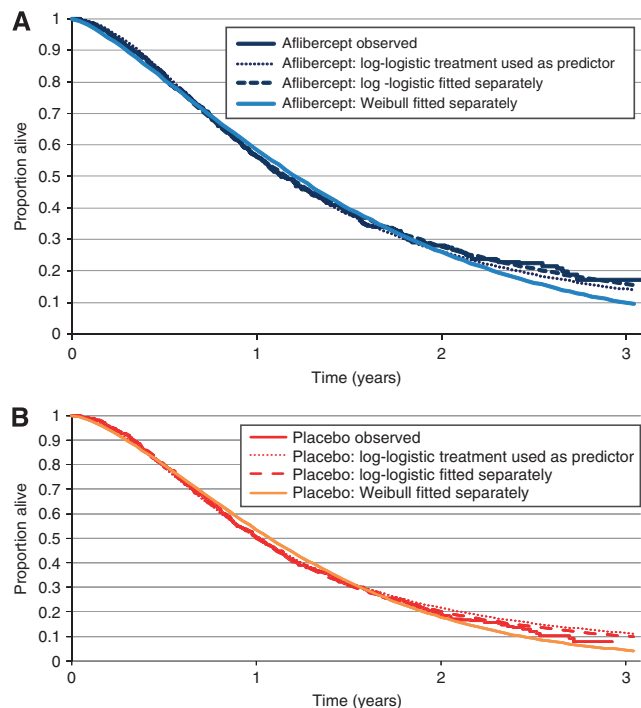

Figure 1. (**A**) Observed *vs* predicted OS (aflibercept ITT Population).
(**B**) Observed *vs* predicted OS (placebo ITT population).

end of the follow-up period, and by the fact that applying the log-logistic distribution for each treatment arm separately provided, visually, a better fit compared with the log-logistic distribution with the treatment arm as predictor (Figure 1A and B). Accordingly, log-logistic distributions fitted for each treatment arm separately provided the best fit to the empirical data. The fit of the log-logistic distribution appeared good visually, as the empirical and predicted curves were almost identical over the observed follow-up period (Figures 1A, B, 2A and B). The Weibull distribution had a worse fit to the observed Kaplan–Meier curve compared with that for the log-logistic distribution, based on both visual inspection and statistical criteria (Table 3, Figures 1A, B, 2A and B). Long-term OS

predictions from the log-logistic *vs* Weibull distributions for the ITT population and pre-defined subgroups are presented in Figures 3 and 4. The log-logistic and the log-normal distributions have a longer tail than that of the Weibull distribution. The 5-year survival rate for the ITT population assuming the log-logistic distribution was 7.2% and 3.9% in aflibercept and placebo arms, respectively, compared with 1.1% and 0.1% if the Weibull distribution was assumed.

Mean overall survival

ITT. Assuming a separately fitted log-logistic distribution, the mean overall survival over 15 years was estimated at 22.8 months for aflibercept + FOLFIRI and at 18.1 months for placebo + FOLFIRI leading to 4.7 months difference (95% CI: 2.1; 6.1). When the Weibull distribution was assumed, the resulting difference in mean OS over 15 years was 3.0 months (1.2; 4.2) (Table 4). Even with the most conservative assumption (i.e., using the Weibull distribution to project survival over the long term), the estimated difference in mean OS was more than double that of the difference of 1.44 months between the observed median survival times of the two treatment arms. When the cutoffs of 5 and 10 years were applied, the mean OS difference using log-logistic distribution was 3.0 and 4.2 months, respectively. Using the Weibull distribution, the difference in mean OS between aflibercept + FOLFIRI and placebo + FOLFIRI was 3.0 months, regardless of the cutoff used for the analyses.

Subgroups. The greatest estimated mean survival gain over 15 years in patients treated with aflibercept + FOLFIRI compared with placebo + FOLFIRI from the log-logistic model was in patients with liver metastasis only (gain of 6.7 months (2.2–10.0)) and in patients with no previous exposure to bevacizumab (6.7 months (3.2–8.1)). In other pre-defined subgroups, estimated survival gain over 15 years for aflibercept + FOLFIRI from the same model ranged from 4.6 months (0.6–7.7) in patients with at most one organ with metastatic involvement to 5.7 months (1.9–7.6) in patients with ECOG PS 0 at baseline. Using the best-fitting distributions, the greatest estimated mean survival gain over 15 years in patients treated with aflibercept + FOLFIRI compared with those treated with placebo + FOLFIRI was in patients with liver metastasis only (gain of 7.6 months (2.4; 13.1)). As expected, results from the Weibull model predicted lower survival gains in all
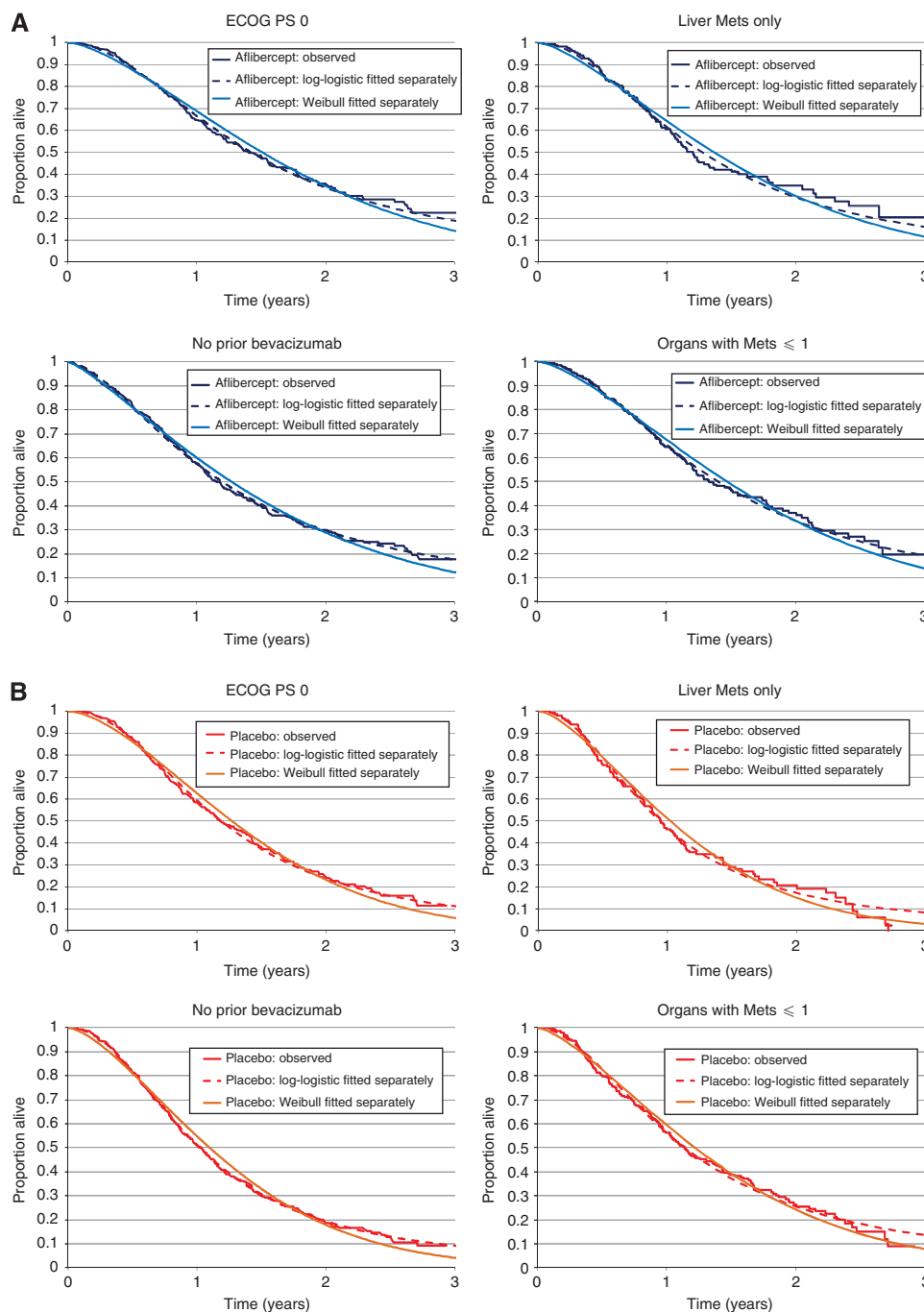
Figure 2. (**A**) Observed *vs* predicted OS for pre-defined subgroups of interest (aflibercept). (**B**) Observed *vs* predicted OS for pre-defined subgroups of interest (placebo).

pre-defined subgroups compared with the log-logistic model, with the smallest difference between treatment groups being 3.4 months (0.5–5.8) in the subgroup of patients with 'number of organs with metastasis $< = 1$' and largest being 5.1 months (1.9–7.9) in patients with liver metastasis only.

For ECOG PS 1 subgroup, estimated survival gain over 15 years from log-logistic model was 2.2 months (-1.1; 5.6) compared with 1.5 months (-0.2; 3.2) from the Weibull model, and, for patients with prior bevacuzimab use, with the log-logistic model it was 0.5 months (-4.4; 5.5) and with the best fit it was 5.1 months (1.2; 9.0) compared with 1.4 months (-0.8; 3.6) (Table 4). It should be noted that for patients with prior

bevacizumab use in the placebo arm, the best fit was provided by the Weibull distribution.

DISCUSSION

The understanding and treatment of mCRC have improved considerably in recent years. Technological advances have led both to better knowledge of tumour biology and appropriate targeting of treatment in metastatic disease, with associated improvements in survival. However, the assessment of the scale

of such benefits may be limited by the dependence on median survival outcomes—in particular OS and PFS. The use of such measures is well-established and likely to continue, as it offers various advantages in cancer intervention trials. For instance, median values are not affected by extreme, unrepresentative results, and thus can give patients and clinicians a sense of typical outcomes achieved with interventions. Inescapably, however, the use of median OS overlooks people whose treatment benefit is such that they remain alive long after study follow-up has finished. Therefore, this approach may limit the assessment of the clinical and health economic consequences of the interventions, analyses that assess mean benefits and mean costs. In our study, we attempted to find a credible and reproducible way of estimating mean OS, given that it is not practical to extend follow-up indefinitely in order to measure this parameter.

The need for such methodology is exemplified by the findings of the VELOUR study. Here, there was a consistent separation of the survival curves in the aflibercept (ziv-aflibercept in United States) and control groups until the end of the follow-up period, with evidence of improved survival in the former group at 18, 24 and

30 months. For example, 2-year survival rates were 28.0% in the aflibercept arm and 18.7% in the control arm. Therefore, median OS may not provide a full assessment of the efficacy of aflibercept, as it cannot fully capture a potentially significant survival benefit beyond the end of the trial.

The current study sought to address this issue by estimating mean OS through parametric extrapolation of the study survival curves. This is now considered a standard approach in cost-effectiveness analyses for new oncolytics (Latimer, 2011). However, such methodology can be fraught with difficulty, primarily related to the considerable uncertainty regarding the potential events after follow-up and, more specifically, around the handling of censored data and the correct approach to parameterisation of survival estimates. Hence, a meaningful and reliable mean estimation of OS has to incorporate multiple statistical approaches and sensitivity analyses to validate the findings. In this study, we applied a systematic process of handling the survival data with multiple approaches coupled with restrictive sensitivity analyses. Analyses truncating OS curves at 5, 10 and 15 years resulted in 3.0, 4.2 and 4.7 months gain in mean OS, respectively, in patients treated with aflibercept + FOLFIRI compared with those treated with placebo + FOLFIRI. Additional analyses included parametric survival fitting, where both arms were analysed together and treatment was used as a predictor in the model. Estimated mean OS gain over 15 years from the above-mentioned analysis was 2.5 months but with a worse fit of the data. These additional analyses help to understand the potential range of the mean OS based on the VELOUR trial and to increase the confidence around the estimates.

Extrapolation of the study survival curve based on the log-logistic distribution (the best-fitting approach) uses a long tail, suggesting that a small proportion of patients survive for a long time. This is in line with recent evidence from several observational studies in mCRC (Sanoff *et al*, 2008; Adam *et al*, 2009; Masi *et al*, 2009). Survival curves published in these studies also had a long tail, suggesting log-logistic or log-normal distributions of survival. Similarly, in other areas of oncology, log-normal and log-logistic distributions have also been shown to be appropriate for long-term extrapolation of OS (Royston, 2001; Chapman *et al*, 2006;
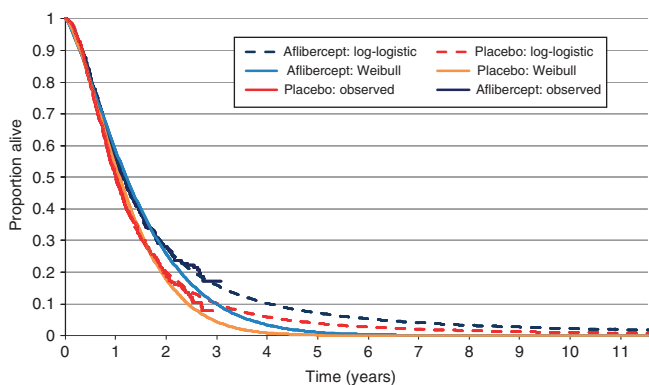

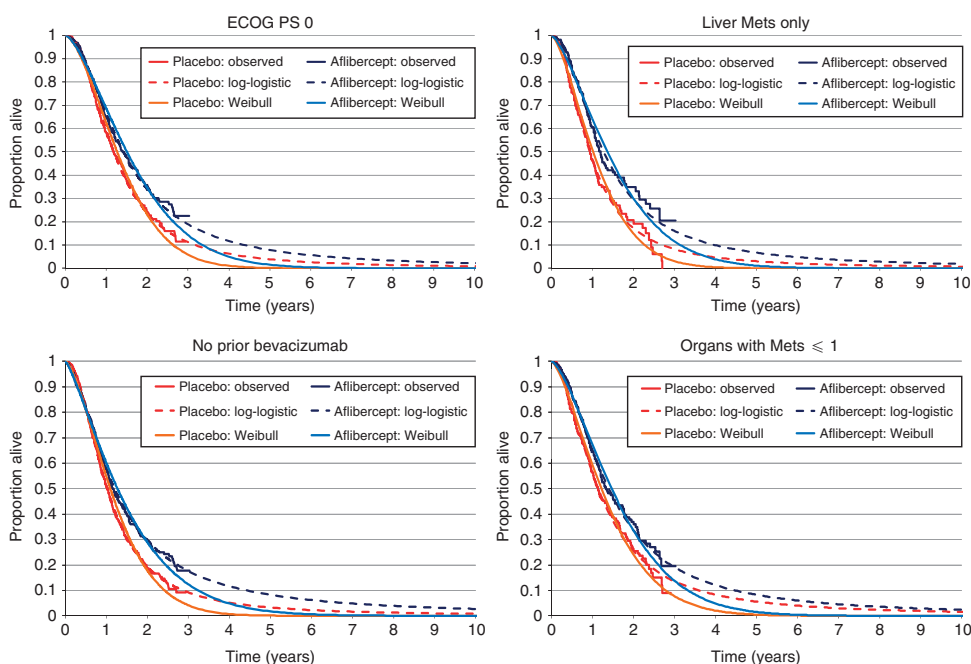Figure 3. Long-term predictions for ITT: log-logistic *vs* Weibull fitted separately.

Figure 4. Long-term predictions for pre-defined subgroups of interest: log-logistic *vs* Weibull fitted separately.

Table 4. Estimated Mean OS (treatment arms fitted separately)			
	Mean OS (months)		
	Placebo/ FOLFIRI	Aflibercept/ FOLFIRI	Difference (95% CI)
ITT			
Best fit: log-logistic	18.9	25.5	6.6 (2.4; 11.6)
Log-logistic with cutoff applied (15 years)	18.1	22.8	4.7 (2.1; 6.1)
Weibull	14.9	17.9	3.0 (1.2; 4.2)
Log-logistic with cutoff applied (5 years)	16.5	19.5	3.0 (1.2; 4.9)
Log-logistic with cutoff applied (10 years)	17.6	21.9	4.2 (1.7; 6.7)
ECOG PS 0			
Best fit: log-normal for placebo/FOLFIRI, log-logistic for aflibercept/FOLFIRI with cutoff applied (15 years)	19.1	25.5	6.4 (2.4; 10.4)
Log-logistic	20.3	27.4	7.1 (2.1;12.8)
Log-logistic with cutoff applied (15 years)	19.8	25.5	5.7 (1.9; 7.6)
Weibull	17.0	20.8	3.8 (1.9; 5.7)
Log-logistic with cutoff applied (5 years)	18.4	22.2	3.8 (1.2; 6.4)
Log-logistic with cutoff applied (10 years)	19.5	24.7	5.2 (1.5; 8.9)
No prior bevacizumab			
Best fit: log-logistic	18.2	28.0	9.8 (4.5; 16.5)
Log-logistic with cutoff applied (15 years)	17.6	24.3	6.7 (3.2; 8.1)
Weibull	15.1	18.8	3.7 (2.0; 5.4)
Log-logistic with cutoff applied (5 years)	16.4	20.3	3.9 (1.8; 6.1)
Log-logistic with cutoff applied (10 years)	17.4	23.2	5.8 (2.7; 8.9)
Liver metastasis only			
Best fit: log-normal for placebo/FOLFIRI, log-logistic for aflibercept/FOLFIRI with cut-off applied (15 years)	15.7	23.3	7.6 (2.4; 13.1)
Log-logistic	17.1	25.0	7.9 (1.2; 16.1)
Log-logistic with cutoff applied (15 years)	16.5	23.3	6.7 (2.2; 10.0)
Weibull	14.1	19.2	5.1 (1.9; 7.9)
Log-logistic with cutoff applied (5 years)	15.4	20.4	5.0 (1.6; 8.6)
Log-logistic with cutoff applied (10 years)	16.3	22.6	6.3 (1.6; 11.3)
Number of organ with metastasis $<=1$			
Best fit: log-normal for placebo/FOLFIRI, log-logistic for aflibercept/FOLFIRI with cutoff applied (15 years)	20.1	25.7	5.7 (0.6; 10.9)
Log-logistic	22.5	28.0	5.5 ($-$1.6; 13.6)
Log-logistic with cutoff applied (15 years)	21.1	25.7	4.6 (0.6; 7.7)
Weibull	17.1	20.5	3.4 (0.5; 8.1)
Log-logistic with cutoff applied (5 years)	18.7	22.1	3.4 (0.3; 6.5)
Log-logistic with cutoff applied (10 years)	20.5	24.8	4.3 (0.0; 8.8)
ECOG PS 1			
Best fit: Log-logistic	15.5	18.3	2.8 ($-$1.5; 7.6)
Log-logistic with cutoff applied (15 years)	14.9	17.1	2.2 ($-$1.1; 5.6)
Weibull	12.3	13.8	1.5 ($-$0.2; 3.2)
Log-logistic with cutoff applied (5 years)	13.8	15.3	1.6 ($-$0.8; 3.9)
Log-logistic with cutoff applied (10 years)	14.6	16.6	2.0 ($-$1.1; 5.2)
Prior bevacizumab use			
Best fit: Weibull for placebo/FOLFIRI, log-logistic for aflibercept/FOLFIRI with cutoff applied (15 years)	14.5	19.5	5.1 (1.2; 9.0)
Loglogistic	20.9	20.7	$-$0.3 ($-$8.2; 6.8)
Log-logistic with cutoff applied (15 years)	19.0	19.5	0.5 ($-$4.4; 5.5)
Weibull	14.5	15.9	1.4 ($-$0.8; 3.6)
Log-logistic with cutoff applied (5 years)	16.6	17.6	1.0 ($-$2.3; 4.2)
Log-logistic with cutoff applied (10 years)	18.4	19.1	0.7 ($-$3.7; 5.0)

Christopherson *et al*, 2008), supporting the hypothesis advocated initially by Boag in the 1940s that a small proportion of 'survivors' in the population can achieve long-term remission with treatment (Boag, 1948, 1949).

With regards to the specific estimation of mean OS, it is particularly notable that for the ITT population, regardless of the distributional assumptions and computational methods, survival advantage in OS was estimated to be at least 3 months for patients treated with aflibercept + FOLFIRI compared with those who received placebo + FOLFIRI.

As with any studies attempting to estimate measures with high uncertainty, our study has associated limitations. In particular,

when distribution models are generated and applied to estimate survival, and regardless of how well they appear to predict outcome, their results are ultimately speculative and uncertain. In that regard, they suffer somewhat by comparison with median survival outcomes, which, despite their limitations, are easily calculated and understood as a stand-alone measure. Furthermore, mean OS calculated by the type of parametric methods we have illustrated must be accompanied by explanation and justification of the approach used. As the choice of an extrapolation technique clearly has the potential to vary from study to study (let alone from one disease to another), this raises obvious questions about the robustness and transferability of the generated data. The use of a single source of data also limits wider interpretation, and the results could be strengthened by further applications of the methodology to mCRC and other cancer intervention trials (Connock et al, 2011). Finally, there is the argument that mean OS itself can be potentially an unrepresentative measure in situations where a very small number of people live for a very long time but the vast majority do not. However, a response to this is that mean and median values of OS both provide useful information on clinical benefits for patients, depending on the specific issues under consideration. Further lines of treatment could also potentially affect the OS estimates. However, in the VELOUR trial, ~32% of patients in each treatment arm (placebo arm 32.1%; aflibercept arm 31.9%) received further treatment with biologics. Cetuximab and bevacizumab (which, as the potentially effective third-line treatments, could have affected OS) had a similar distribution between the treatment arms (9.0% and 12.2% of patients received bevacizumab in the aflibercept and placebo arms, respectively; 17.6% and 14.8% of patients received cetuximab in the aflibercept and placebo arms, respectively).

In summary, the mean overall survival gains from aflibercept in patients with previously treated mCRC in the VELOUR trial were estimated to be at least 3 months (using a 5-year cutoff). We believe that our study presents a robust and comprehensive estimation of the added benefit with respect to mean OS. These findings provide payers, health economists and clinicians with a broader evidence base for assessing potential benefits of the drug and may have important implications for both clinical and economic decision-making in mCRC.

CONFLICT OF INTEREST

The following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a 'U' are those for which no compensation was received; those relationships marked with a 'C' were compensated. Employment or Leadership Position: Florence Joulain (C) and Sheikh Usman Iqbal (C) are employees of Sanofi. Consultant or Advisory Role: Josep Tabernero, Sanofi (C), Regeneron Pharmaceuticals (C); Carmen Allegra, Sanofi (C). Stock Ownership: Florence Joulain, Sanofi; Sheikh Usman Iqbal, Sanofi. Honoraria: Josep Tabernero, Sanofi. Sanofi Research Funding: Eric Van Cutsem, Sanofi. Martin Hoyle is an employee of the University of Exeter, UK. He received honoraria from United BioSource Corporation to advise on the research described in this article. Irina Proskorovsky is employed by Evidera, which provides consulting and other research services to pharmaceutical, device, government and non-government organisations, and received funding from Sanofi to conduct the research described in this article. In this salaried position, Irina Proskorovsky works with a variety of companies and organisations. She receives no payment or honoraria directly from these organisations for services rendered.

REFERENCES

Adam R, Wicherts DA, de Haas RJ, Ciacio O, Levi F, Paule B, Ducreux M, Azoulay D, Bismuth H, Castaing D (2009) Patients with initially unresectable colorectal liver metastases: is there a possibility of cure? *J Clin Oncol* **27**(11): 1829–1835.

Boag J (1948) The presentation and analysis of the result of radiotherapy. *Br J Radiol* **21**(244): 189 passim.

Boag J (1949) Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J Roy Stat Soc Ser B (Methodological)* **11**: 15–53.

Chapman JA, Lickley HL, Trudeau ME, Hanna WM, Kahn HJ, Murray D, Sawka CA, Mobbs BG, McCready DR, Pritchard KI (2006) Ascertaining prognosis for breast cancer in node-negative patients with innovative survival analysis. *Breast J* **12**(1): 37–47.

Christopherson R, James KE, Tableman M, Marshall P, Johnson FE (2008) Long-term survival after colon cancer surgery: a variation associated with choice of anesthesia. *Anesth Analg* **107**(1): 325–332.

Cleves M, Gould W, Gutierrez R, Marchenko Y (2002) *An Introduction to Survival Analysis using Stata*. Stata Press: College Station, TX.

Connock M, Hyde C, Moore D (2011) Cautions regarding the fitting and interpretation of survival curves: examples from NICE single technology appraisals of drugs for cancer. *Pharmacoeconomics* **29**(10): 827–837.

Davies A, Briggs A, Schneider J, Levy A, Ebeid O, Wagner S, Kotapati S, Ramsey S (2012) The Ends Justify the Mean: Outcome Measures for Estimating the Value of New Cancer Therapies. *Health Outcomes Research in Medicine* **3**(1): e25–e36.

Davies JM, Goldberg RM (2008) First-line therapeutic strategies in metastatic colorectal cancer. *Oncology (Williston Park)* **22**(13): 1470–1479.

Goldberg RM, Rothenberg ML, Van Cutsem E, Benson 3rd AB, Blanke CD, Diasio RB, Grothey A, Lenz HJ, Meropol NJ, Ramanathan RK, Becerra CH, Wickham R, Armstrong D, Viele C (2007) The continuum of care: a paradigm for the management of metastatic colorectal cancer. *Oncologist* **12**(1): 38–50.

Hess GP, Wang PF, Quach D, Barber B, Zhao Z (2010) Systemic therapy for metastatic colorectal cancer: patterns of chemotherapy and biologic therapy use in US medical oncology practice. *J Oncol Pract* **6**(6): 301–307.

Hind D, Tappenden P, Tumur I (2008) The use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer: systematic review and economic evaluation. *Health Technol Assess* **12**: 15.

Huse DM, von Mehren M, Lenhart G, Joensuu H, Blanke C, Feng W, Finkelstein S, Demetri G (2007) Cost effectiveness of imatinib mesylate in the treatment of advanced gastrointestinal stromal tumours. *Clin Drug Invest* **27**(2): 85–93.

International Agency for Research on Cancer. WHO, Cancer Research UK (2011) CancerStats—Cancer Worldwide.

Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. *CA Cancer J Clin* **61**(2): 69–90.

Latimer N (2011) NICE DSU Technical Support. Undertaking survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data. Available at http://www.nicedsu.org.uk/ NICE%20DSU%20TSD%20Survival%20analysis_finalv2.pdf. Accessed on 25 June 2012.

Masi G, Loupakis F, Pollina L, Vasile E, Cupini S, Ricci S, Brunetti IM, Ferraldeschi R, Naso G, Filipponi F, Pietrabissa A, Goletti O, Baldi G, Fornaro L, Andreuccetti M, Falcone A (2009) Long-term outcome of initially unresectable metastatic colorectal cancer patients treated with 5-fluorouracil/leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) followed by radical surgery of metastases. *Ann Surg* **249**(3): 420–425.

Muszbek N, Kreif N, Valderrama A, Agnes B, Ishak J, Ross P (2012) Modeling survival in hepatocellular carcinoma. *Curr Med Res Opin* **28**(7): 1141–1153.

Neymark N, Adriaenssen I, Gorlia T, Caleo S, Bolla M (2002) Estimating survival gain for economic evaluations with survival time as principal endpoint: a cost-effectiveness analysis of adding early hormonal therapy to

radiotherapy in patients with locally advanced prostate cancer. *Health Econ* **11**: 233–248.

Rosé M, Iqbal SU, Naoshy S, Mehta J, Chau I (2012) Current Treatment Patterns in First-Line (1L) and Second-Line (2L) Treatment for Metastatic Colorectal Cancer (mCRC) in Western Europe. *Accepted and poster presented at ESMO World Congress GI June 2012.*

Royston P (2001) The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Stat Neerl* **2**: 89–104.

Sanoff HK, Sargent DJ, Campbell ME, Morton RF, Fuchs CS, Ramanathan RK, Williamson SK, Findlay BP, Pitot HC, Goldberg RM (2008) Five-year data and prognostic factor analysis of oxaliplatin and irinotecan combinations for advanced colorectal cancer: N9741. *J Clin Oncol* **26**(35): 5721–5727.

Shilov GEe, Gurevich BL, Silverman RA (1977) *Integral, Measure and Derivative: a Unified Approach*, Richard A Silverman (edn). New York: Dover Publications: London: Constable.

Singer J, Willett J (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press: Oxford.

Tournigand C, Andre T, Achille E, Lledo G, Flesh M, Mery-Mignard D (2004) FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. *J Clin Oncol* **22**(2): 229–237.

Van Cutsem E, Tabernero J, Lakomy R, Prenen H, Prausová J, Macarulla T, Ruff P, van Hazel GA, Moiseyenko V, Ferry D, McKendrick J, Polikoff J, Tellier A, Castan R, Allegra C (2012) Addition of aflibercept to folfiri improves survival in a phase iii randomized trial in patients with metastatic colorectal cancer previously treated with an oxaliplatin-based regimen. *J Clin Oncol* **30**(28): 3499–3506.