

Projecting Event-Based Analysis Dates in Clinical Trials: An Illustration Based on the International Duration Evaluation of Adjuvant Chemotherapy (IDEA) Collaboration. Projecting analysis dates for the IDEA collaboration

Lindsay A. Renfro*¹, Axel M. Grothey², James Paul³, Irene Floriani⁴, Franck Bonnetain⁵, Donna Niedzwiecki⁶, Takeharu Yamanaka⁷, Ioannis Souglakos⁸, Greg Yothers⁹, Daniel J. Sargent¹

¹*Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA*

²*Department of Oncology, Mayo Clinic, Rochester, MN, USA*

³*Cancer Research UK, Beatson West of Scotland Cancer Centre, Institute of Cancer Sciences, University of Glasgow, Glasgow, Scotland*

⁴*Department of Oncology, IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy*

⁵*Methodological and Quality of Life Unit in Oncology, University Hospital, Besançon, France*

⁶*Department of Biostatistics, Duke University, Durham, NC, USA*

⁷*Department of Biostatistics, National Cancer Center, Tokyo, Japan*

⁸*Division of Internal Medicine, University of Crete, Crete, Greece*

⁹*National Surgical Adjuvant Breast and Bowel Project Biostatistical Center, Pittsburgh, PA, USA*

Abstract: Purpose: Clinical trials are expensive and lengthy, where success of a given trial depends on observing a prospectively defined number of patient events required to answer the clinical question. The point at which this analysis time occurs depends on both patient accrual and primary event rates, which typically vary throughout the trial's duration. We demonstrate real-time analysis date projections using data from a collection of six clinical trials that are part of the IDEA collaboration, an international preplanned pooling of data from six trials testing the duration of adjuvant chemotherapy in stage III colon cancer, and we additionally consider the hypothetical impact of one trial's early termination of follow-up.

Patients and Methods: In the absence of outcome data from IDEA, monthly accrual rates for each of the six IDEA trials were used to project subsequent trial-specific accrual, while historical data from similar Adjuvant Colon Cancer Endpoints (ACCENT) Group trials were used to construct a parametric model for IDEA's primary endpoint, disease-free survival, under the same treatment regimen. With this information and using the planned total accrual from each IDEA trial protocol, individual patient accrual and event dates were simulated and the overall IDEA interim and final analysis times projected. Projections were then compared with actual (previously undisclosed) trial-specific event totals at a recent census time for validation. The change in projected final analysis date assuming early termination of follow-up for one IDEA trial was also calculated.

Results: Trial-specific predicted event totals were close to the actual number of events per trial for the recent census date at which the number of events per trial was known, with the overall IDEA projected number of events only off by eight patients. Potential early termination of follow-up by one IDEA trial was estimated to postpone the overall IDEA final analysis date by 9 months.

Conclusions: Real-time projection of the final analysis time during a trial, or the overall analysis time during a trial collaborative such as IDEA, has practical implications for trial feasibility when these projections are translated into additional time and resources required.

Keywords: *Colon cancer • Adjuvant therapy • Meta-analysis • Duration of therapy*

© De Gruyter Open

* E-mail: renfro.lindsay@mayo.edu

Introduction

Clinical trials are expensive and lengthy endeavours, with success of a given trial dependent upon collection of a sufficient amount of primary endpoint information to draw meaningful trial conclusions (for the desired statistical power), whether positive or negative. In randomized trials with time-to-event primary endpoints, the amount of such information is prospectively defined in the protocol in terms of the number of events, or patient outcomes, that must be observed before an analysis (interim or final) can be performed. The required number of observed outcomes is a function of the desired effect size, power and type I error rates for the trial, the expected rate of patient accrual and the distribution of the time-to-event endpoint itself (usually summarized by a median event time, assuming the underlying time-to-event outcome is exponentially distributed). Once a trial opens to enrolment and patient follow-up and treatment begins, however, the accrual rate may fluctuate over time, and the distribution of the time-to-event outcome may not closely resemble an exponential distribution. While the latter issue can be prevented through careful modelling of similar data from historical trials during initial planning of the study, accrual issues commonly plague already-launched clinical trials in practice, leading the study team and funding entities to question a moving target: the time remaining and additional resources required to reach the amount of information (events) necessary to perform interim and final analyses.

A number of existing works have addressed related issues such as mid-trial sample-size recalculation or adaptive use of real-time endpoint information to make termination or continuation decisions; reviews of these topics are given by Friede and Kieser (2006) and Kairalla et al. (2012), respectively.¹⁻² A more common situation arises when mid-trial changes to the design are not necessarily desired, but the remaining trial duration (in months or years) would be useful to estimate. In these cases, generally all that is readily known during the study is the accrual rate (per some period of time, such as months) since the study opened, the number of events already observed (unless some parties are blinded to this information) and some external estimate or hypothesis of the distribution of event times for the specific disease and treatment(s) under study. In some cases, it is also learned mid-trial that patient outcomes observed thus far are better than what was expected during trial planning, such that adaptations to a null hypothesized event rate must be considered. Bagiella and Heitjan (2001) proposed two methods for forecasting analysis

dates: one that extrapolates cumulative mortality into the future until the number of required events (deaths) for the analysis is reached and another that produces a Bayesian prediction based on a predictive distribution of milestone event times.³ However, both of these methods are fully parametric, assuming exponentially distributed event times and enrolment times following a Poisson distribution; because of this, deviations from underlying distributional assumptions could easily result in bias or inefficiency. To address these issues, Ying et al. (2004) proposed a nonparametric method for the prediction of event times in randomized clinical trials, which extrapolates Kaplan–Meier probabilities into the future using a Bayesian bootstrap to identify the time at which the required number of events for an analysis may occur.⁴ These authors showed that their nonparametric approach is superior when the underlying time-to-event outcome violates the assumptions of Bagiella and Heitjan, yet their method is informed only by data from the trial itself and does not make use of existing knowledge (e.g. similar historical data) that is often available and could improve endpoint modelling and thus analysis time predictions in the present trial. Furthermore, these and other existing methods rely heavily on rather sophisticated (e.g. Bayesian) statistical techniques, which may present practical limitations for many trial practitioners. In this paper, we demonstrate and validate an algorithm for obtaining predictions of interim and final analysis times, through application of relatively straightforward parametric and non-parametric methodologies to a multi-trial collaboration in early-stage colon cancer.

Motivating Example: IDEA Collaboration

The work presented here was motivated by a specific example: the International Duration Evaluation of Adjuvant Therapy (IDEA) Collaboration, a collection of six concurrent clinical trials whose joint mission is to test the non-inferiority of 3 months of oxaliplatin-based adjuvant chemotherapy for the treatment of stage III colon cancer versus the standard 6 months' administration of the same therapy, with the primary endpoint of disease-free survival (DFS). In this setting, DFS is defined as the time from study randomization to disease recurrence or death, whichever occurs earlier. Full details of the individual IDEA trials and rules governing the collaboration, including data sharing, have been described by André et al. (2013).⁵ According to the IDEA statistical analysis plan, 3 months of adjuvant therapy will be deemed non-inferior to 6 months of therapy if the two-sided 95% confidence interval for

Table 1. IDEA Trial Characteristics and Status of Trials as of 31 January 2014.

Trial	Primary Country	Group	Month First Patient Enrolled	January 2014 Status	Planned Stage III Accrual	Stage III Accrual by 31 January 2014
SCOT	UK	---	July 2008	Closed	4000	4015
TOSCA	Italy	GISCAD	July 2007	Closed	2500	2444
PRODIGE	France	GERCOR/PRODIGE	May 2009	Open	2000	1882
C80702	USA	CALGB/SWOG	July 2010	Open	2500	1536
HORG	Greece	HORG	Sept 2010	Open	1000	560
ACHIEVE	Japan	JFMC	Aug 2012	Open	1200	992
TOTAL	---	IDEA	July 2007	Open	12,500	11,429

SCOT, Short Course Oncology Therapy; TOSCA, Three or Six Colon Adjuvant; CALGB, Cancer and Leukaemia Group B, SWOG, Southwest Oncology Group; HORG, Hellenic Oncology Research Group; JFMC, Japanese Foundation for Multidisciplinary Treatment of Cancer; IDEA, International Duration Evaluation of Adjuvant Chemotherapy.

the hazard ratio lies entirely below 1.12. A total sample size of at least 10,500 patients across the six trials (with a targeted 12,500 patients if each trial fully enrolls) is required to observe 3,390 DFS events to achieve 90% power for this design, assuming a 3-year DFS rate of 0.72 in the control group. The individual studies opened to enrolment between 2007 and 2012, and as of January 2014 (when the predictions reported in this paper were requested), two studies had completed their planned accrual while the other four studies remained open for enrolment (Table 1).

Patients and Methods

Predicted interim and final analysis times for the IDEA trial collaboration were derived from the actual monthly accrual rates supplied by each trial, in combination with simulated patient outcomes, where features of the latter were estimated from similarly treated stage III patients enrolled to previously completed clinical trials contained in the Adjuvant Colon Cancer Endpoints (ACCENT) database.⁶⁻¹¹ The ACCENT database is a collection of patient-level data from more than 40,000 patients enrolled to randomized clinical trials for adjuvant treatment of early-stage colon cancer since 1977. For the purposes of this analysis, only ACCENT patients from recent trials with oxaliplatin-containing arms were considered for endpoint estimation, as detailed below.

IDEA Accrual Determination

Historical monthly accrual of stage III patients was reported by each IDEA clinical trial, and these accrual rates plotted over time for visual examination. After increasing or decreasing trends in recent months'

accrual were ruled out, future monthly accrual rates for each of the four IDEA trials still enrolling patients were projected by carrying forward the average monthly accrual for the past 12 months until the protocol-defined maximum accrual for stage III patients was met.

ACCENT Outcome Estimation

Separately, possible parametric models for DFS were examined using historical outcomes from patients who received 6 months planned FOLFOX (FLOX in C-07) or XELOX (the control arms in IDEA) while enrolled on the trials C-07⁸, C-08⁹, N016968 (XELOXA)¹⁰ and N0147¹¹ contained in the ACCENT database. Patients from the FOLFOX + Bevacizumab arm of C-08 and the FOLFOX + Cetuximab arm of N0147 were included as each of those individual trials showed no improvement from the addition of the experimental therapy. Each of these trials used DFS as the primary endpoint and enrolled patients within the past 15 years, such that the model with the best fit to these 6,537 patients could subsequently be used to simulate hypothetical outcomes for similarly treated IDEA patients. Candidate parametric models included the Weibull, Gompertz, Log-Normal, Log-Logistic, Generalized Gamma and Generalized F distributions (see, e.g. Kalbfleish and Prentice (2002)¹¹ and Prentice (1975)¹²). Goodness of fit was determined visually by superimposing the fitted model from each family onto a Kaplan–Meier curve of the ACCENT patients and checking for best proximity over all time points. Formal analytical goodness-of-fit testing was planned for cases where visual comparisons would not have yielded a clear winner¹². Rates of patient loss to follow-up over time were also examined in ACCENT, such that simulated outcomes for IDEA patients could be modified to account for similar potential loss to follow-up.

IDEA Patient Simulation

Using both the monthly trial-specific accrual from IDEA and the parametric estimation of DFS from ACCENT, hypothetical patient outcomes were simulated for each trial within the IDEA collaboration as follows. First, random accrual dates were generated from the reported and projected monthly accrual rates, assuming uniform accrual within each trial and month. Next, a random DFS time for each patient was simulated from the best-fitting parametric model previously estimated from ACCENT. These simulated patient DFS times were then added to the simulated accrual dates to produce a simulated event date for each patient. Early patient dropout was accounted for by randomly and uniformly right-censoring IDEA patients at the same rates estimated within monthly intervals from ACCENT. This process of IDEA patient data simulation described above was repeated over 100 iterations, such that stable analysis time predictions based on averages taken over hypothetical “runs” of the IDEA collaboration could be obtained. In this setting, 100 iterations were deemed sufficient as the variability across iterations was low and computational time was non-negligible.

Interim and Final Analysis Time Predictions

After accrual and event times were simulated for all IDEA patients, the predicted interim and final analysis dates were derived by calculating the cumulative number of predicted events by month to identify those months where 1695 and 3390 events were reached, respectively, and then averaging these predictions over the 100 simulation iterations. The hypothetical impact on the final IDEA analysis date of the SCOT trial’s possible discontinuation of follow-up beyond November 2014 was also considered. Importantly, the actual numbers of events per trial at past quarterly census dates were withheld by the IDEA steering committee from the statistician performing this analysis until *after* the predictions were completed. At that time, the most recently reported number of events per trial was compared with the trial-specific event predictions from the same month, assuming a 4-month “information lag time” between clinical documentation of patient events and submission of these events to centralized trial databases. This decision readily allowed for evaluation and validation of the prediction methodology.

Results

IDEA Accrual Determination

Monthly accrual for each trial since July 2007, when the first trial (TOSCA) opened to enrolment, is shown in

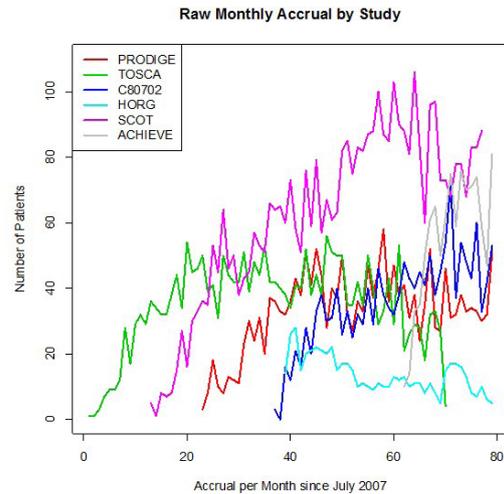


Figure 1. Monthly accrual by trial.

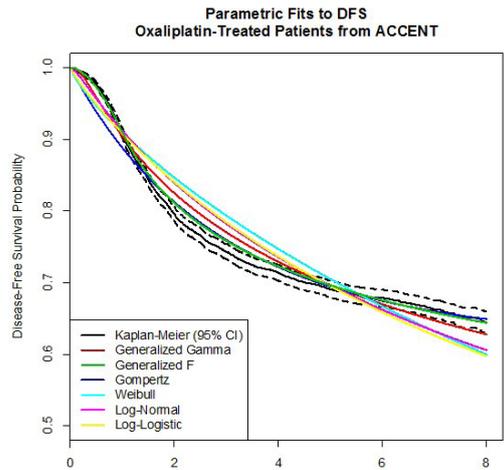


Figure 2. Parametric fits to ACCENT data.

Figure 1. From these patterns, it was determined that future monthly accrual rates for the four IDEA trials that remained open as of January 2014 (PRODIGE, C80702, HORG and ACHIEVE) could reasonably be estimated by their respective monthly accrual averages calculated over the past 12 months.

ACCENT Outcome Estimation

In Figure 2, the candidate parametric models for DFS applied to the historically similar ACCENT patients are shown, superimposed on the empirical Kaplan–Meier estimate for DFS in ACCENT with its 95% confidence band. Among these candidate distributions, it was visually determined that the Generalized F distribution yielded the best parametric fit to the empirical data, as represented by its general proximity to the Kaplan–

Meier curve at most follow-up times. Details of the generalized F distribution and its parameter estimates from ACCENT are given in the Appendix. Because of this clear winner, no additional numerical goodness-of-fit testing was performed. Further analysis of the ACCENT data revealed monthly loss to follow-up rates ranging from less than 0.1% prior to 5 years to more than 1.5% beyond 5 years, which were the follow-up loss rates applied to the simulated patient outcomes in IDEA.

Interim and Final Analysis Time Predictions

With DFS event times simulated for each patient in IDEA, the cumulative number of predicted events was computed, both within trials and overall, for each calendar month since July 2007. Figure 3 presents the averages of these predictions taken over the 100 simulation iterations. Based on these predictions, it was estimated that the 1,695th event required to perform

the single pre-specified interim analysis would occur in January 2014 (simulated IQR: December 2013 to January 2014), and the 3,390th event required to perform the final analysis (if the interim analysis allowed the project to continue) would occur in September 2016 (simulated IQR: August 2016 to October 2016). For the hypothetical scenario in which the SCOT trial discontinued patient follow-up beyond November 2014, it was predicted that the 3,390th event could not be observed until June 2017, causing a final analysis delay of approximately 9 months (Figure 4).

The total event projections (lines) shown in Figures 3 and 4 were constructed without knowledge of the actual number of events at any point during the IDEA collaboration, such that the predictions could not be biased by “back fitting” the models to known quantities. Once analysis date predictions were provided to members of the IDEA steering committee, the number

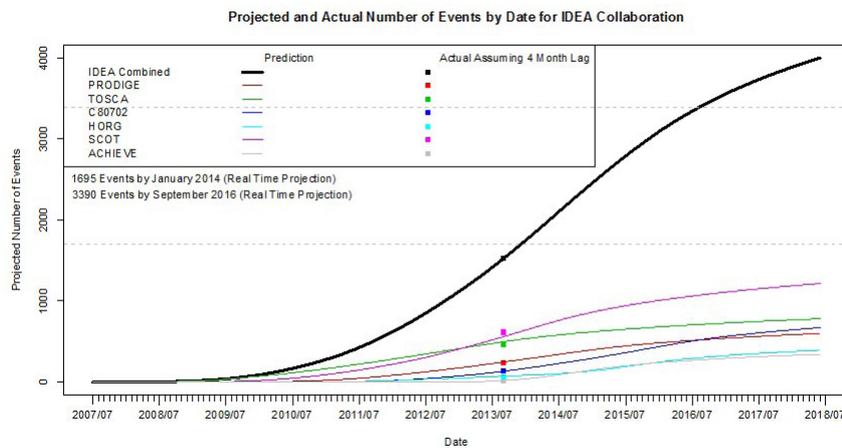


Figure 3. IDEA event and analysis time projections: overall.

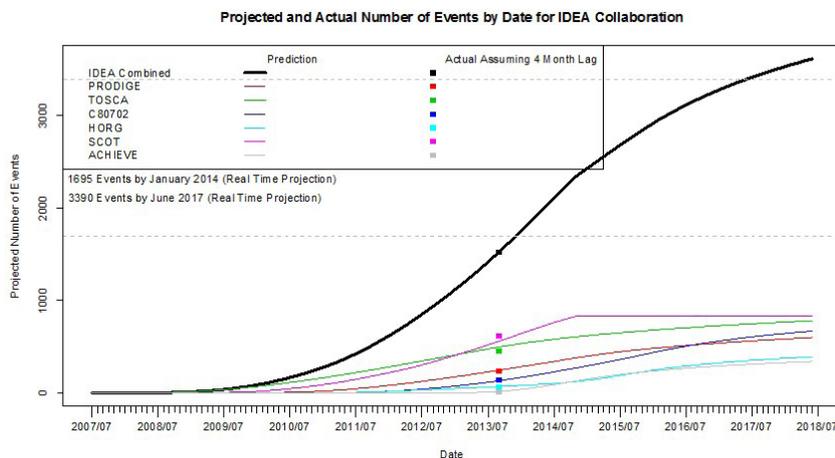


Figure 4. IDEA event and analysis time projections: without SCOT trial.

of events per trial and the total number of events overall as of the last quarterly census (January 2014) were disclosed. These counts were then added to Figures 3 and 4 (represented by solid squares), assuming a 4-month information lag time between the actual occurrence of events and when they are centrally recorded in trial databases. The high concordance observed between the predicted and actual number of events, both within trials (a difference of fewer than 10 events for 4 out of 6 trials) and overall (a difference of only 8 events), suggests a strong resemblance of IDEA patients to those in ACCENT who received oxaliplatin-containing regimens, as well as a strong parametric fit of the generalized F distribution to the ACCENT data.

Discussion

In this manuscript, using the IDEA trial collaboration as an example, we have demonstrated a useful methodology for estimating the eventual analysis time in a clinical trial where a time-to-event outcome is of interest. Having a reasonably accurate idea of the final analysis date well in advance allows for trial resource allocation and planning to occur more efficiently and offers an opportunity for foresight in situations where it becomes evident that adjustments to trial objectives or design characteristics should be made. As we have shown, the eventual analysis time in a clinical trial is a function of patient accrual, the distribution of the patient outcome being studied and consistency of follow-up. As was shown in the hypothetical scenario where one IDEA trial terminated follow-up prematurely, losing even a fraction of the endpoint information can profoundly impact the time until a final analysis can be performed. In single trial settings, this can occur when individual centers within a trial fail to report patient data in a timely manner, or when a higher number of patients than anticipated are lost to follow-up. Even as trial characteristics such as the amount of patient information (number of events) required for the final analysis are generally determined well before the first patient is enrolled, utilization of accumulating endpoint data—in a neutral and prospective manner

such as that presented here—may be useful for judging whether a primary trial objective is still on track to be achievable within a reasonable amount of time, or revealing the necessity for design changes while they can still be made.

Acknowledgements

The authors acknowledge the contributions made by the principal investigators of each of the IDEA trials: Timothy Iveson, M.D. (SCOT); Roberto Labianca, M.D. and Alberto Sobrero, M.D. (TOSCA); Jeffrey Meyerhardt, M.D., M.P.H. and Anthony Shields, M.D., Ph.D. (C80702); Thierry André, M.D. and Julien Taieb, M.D. (PRODIGE-GERCOR); Atsushi Ohtsu M.D., and Takayuki Yoshino, M.D. (ACHIEVE); Ioannis Souglakos, M.D. (HORG).

Appendix

The generalized F distribution is derived as follows. If $y \sim F(2s_1, 2s_2)$ and $w = \log(y)$, then $x = \exp(w\sigma + \mu)$ has a generalized F distribution with location parameter μ , scale parameter $\sigma > 0$, and shape parameters s_1 and s_2 . In a more stable version described by Prentice¹², s_1 and s_2 are replaced by shape parameters Q and $P > 0$, where $s_1 = 2/(Q^2 + 2P + Q\delta)$ and $s_2 = 2/(Q^2 + 2P - Q\delta)$. Equivalently, $P = 2/(s_1 + s_2)$ and $Q = (1/s_1 + 1/s_2)/\{(1/s_1 + 1/s_2)^{-1/2}\}$. If we define $\delta = (Q^2 + 2P)^{1/2}$ and $w = \delta\{\log(x) - \mu\}/\sigma$, then the probability density function of x is given by

$$f(x) = \frac{\delta \left(\frac{s_1}{s_2}\right)^{s_1} \exp(s_1 w)}{\sigma x \left\{1 + s_2 \exp\left(\frac{w}{s_2}\right)\right\}^{s_1 + s_2} B(s_1, s_2)}$$

Where $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$ for $a, b > 0$ is the beta function. When DFS among stage III ACCENT patients treated with oxaliplatin-containing arms is assumed to follow a generalized F distribution, the resulting parameter estimates (subsequently used to simulate the IDEA patients) are given by $\hat{\mu} = 2.586$, $\hat{\sigma} = 0.518$, $\hat{P} = 26.14$, and $\hat{Q} = -11.59$.

References

- [1] Friede T, Kieser M. Sample size recalculation in internal pilot study designs: a review. *Biom J* 48: 537-55, 2006
- [2] Kairalla JA, Coffey CS, Thomann MA, Muller KE. Adaptive trial designs: a review of barriers and opportunities. *Trials* 13:145, 2012
- [3] Bagiella E, Heitjan DF. Predicting analysis times in randomized clinical trials. *Stat Med* 20: 2055-63, 2001
- [4] Ying G, Heitjan DF, Chen T. Nonparametric prediction of event times in randomized clinical trials. *Clin Trials* 1:352-261, 2004
- [5] André T, Iveson T, Labianca R, et al. for the IDEA Steering Committee. The IDEA (International

- Duration Evaluation of Adjuvant Chemotherapy) Collaboration: Prospective Combined Analysis of Phase III Trials Investigating Duration of Adjuvant Therapy with the FOLFOX (FOLFOX4 or Modified FOLFOX6) or XELOX (3 versus 6 months) Regimen for Patients with Stage III Colon Cancer: Trial Design and Current Status. *Curr Colorectal Cancer Rep* 9: 261-9, 2013
- [6] Sargent DJ, Wieand HS, Haller DG, et al: Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *J ClinOncol* 23:8664-70, 2005
- [7] Sargent DJ, Shi Q, Yothers G, et al: Two or three year disease free survival (DFS) as a primary endpoint in stage III adjuvant colon cancer trials with fluoropyrimidines with or without oxaliplatin or irinotecan: data from 12,676 patients from MOSAIC, X-ACT, PETACC-3, C-06, C-07, and C89803. *Eur J Cancer* 47:990-6, 2011
- [8] Yothers G, O'Connell M, Allegra CJ, et al. Oxaliplatin as adjuvant therapy for colon cancer: updated results of NSABP C-07 trial, including survival and subset analyses. *J ClinOncol* 28: 3768-74, 2011
- [9] Allegra CJ, Yothers G, O'Connell MJ, et al. Phase III trial assessing bevacizumab in stages II and III carcinoma of the colon: results of the NSABP protocol C-08. *J ClinOncol* 29:11-6, 2011
- [10] Schmoll H-J, Cartwright T, Tabernero J, et al. Phase III trial of capecitabine plus oxaliplatin as adjuvant therapy for stage III colon cancer: a planned safety analysis in 1,864 patients. *J ClinOncol* 25:102-9, 2006
- [11] Alberts SR, Sargent DJ, Smyrk TC, et al. Adjuvant mFOLFOX6 with or without cetuximab in KRAS wild-type patients with resected stage III colon cancer: results from NCCTG intergroup phase III trial N0147. *J ClinOncol* 28: (18_suppl CRA3507), 2010
- [12] Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*, 2nd edition. Hoboken, NJ: John Wiley & Sons, Inc., 2002
- [13] Prentice RL. Discrimination among some parametric models. *Biometrika* 62:607-614, 197