European Journal of Cancer 123 (2019) 101–111

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Disease-free survival as a surrogate for overall survival in neoadjuvant trials of gastroesophageal adenocarcinoma: Pooled analysis of individual patient data from randomised controlled trials☆

Ulrich Ronellenfitsch [a,*], Katrin Jensen [b], Svenja Seide [b], Meinhard Kieser [b], Matthias Schwarzbach [c], Tracy E. Slanger [d], Bryan Burmeister [e], David Kelsen [f], Donna Niedzwiecki [g], Guillaume Piessen [h], Christoph Schuhmacher [i], Susan Urba [j], Cornelis van de Velde [k], Marc Ychou [l], Ralf Hofheinz [m], Sylvie Lorenzen [n]

[a] *Medical Faculty of the Martin Luther University Halle-Wittenberg and University Hospital Halle (Saale), Department of Visceral, Vascular and Endocrine Surgery, Ernst-Grube-Str. 40, 06120, Halle (Saale), Germany*
[b] *Institute of Medical Biometry and Informatics, University of Heidelberg, Im Neuenheimer Feld 130.3, 69120, Heidelberg, Germany*
[c] *Department of General, Visceral, Vascular, and Thoracic Surgery, Klinikum Frankfurt Höchst, Gotenstraße 6-8, 65929, Frankfurt am Main, Germany*
[d] *Institute for Quality and Efficiency in Health Care (IQWiG), Im Mediapark 8, 50670, Cologne, Germany*
[e] *University of Queensland, Princess Alexandra Hospital, Brisbane, Australia, Princess Alexandra Hospital, Brisbane, QLD, 4102, Australia*
[f] *Memorial Sloan-Kettering Cancer Center and Weill Cornell Medical College, 1275 York Avenue, New York, NY, 10065, USA*
[g] *The Alliance for Clinical Trials in Oncology (Alliance) Statistics and Data Center, Duke University Medical Center, Hock Plaza, 2424 Erwin Rd, Room 8040, Durham, NC, 27705, USA*
[h] *Department of Digestive and Oncological Surgery, University Hospital C. Huriez, Inserm UMR-S 1172, Jean-Pierre Aubert Research Center (JPARC) Team, 1, Rue Polonovski, 59037, Lille Cedex, France*
[i] *Department of Surgery, Klinikum Rechts der Isar, Technical University Munich, Ismaninger Str. 22, 81675, Munich, Germany*
[j] *Division of Hematology/Oncology, University of Michigan Medical Center, 1500 E Medical Center Drive, C347, SPC 5848 Ann Arbor, MI, 48109, USA*
[k] *Department of Surgery, Leiden University Medical Center, K6-R, P.O. Box 9600, 2300, RC Leiden, the Netherlands*
[l] *Centre Régional de Lutte Contre le Cancer, Val D'Aurelle, Montpellier Cedex 05, France*

☆ Parts of this work were supported by the German Federal Ministry of Education and Research (grant number 01 KG 0807) and intramural research funds of the University of Heidelberg (Foundation for Cancer Research).

* *Corresponding author*: Fax: +49 345 557-2823
E-mail addresses: ulrich.ronellenfitsch@uk-halle.de (U. Ronellenfitsch), jensen@imbi.uni-heidelberg.de (K. Jensen), seide@imbi.uni-heidelberg.de (S. Seide), meinhard.kieser@imbi.uni-heidelberg (M. Kieser), matthias.schwarzbach@klinikumfrankfurt.de (M. Schwarzbach), t.slanger@hotmail.com (T.E. Slanger), bryan.burmeister@health.qld.gov.au (B. Burmeister), kelsend@mskcc.org (D. Kelsen), donna.niedzwiecki@duke.edu (D. Niedzwiecki), guillaume.piessen@chru-lille.fr (G. Piessen), christoph.schuhmacher@tum.de (C. Schuhmacher), surba@med.umich.edu (S. Urba), c.j.h.van_de_velde@lumc.nl (C. van de Velde), marc.ychou@icm.unicancer.fr (M. Ychou), ralf.hofheinz@umm.de (R. Hofheinz), sylvielorenzen@gmx.de (S. Lorenzen).

https://doi.org/10.1016/j.ejca.2019.10.001

[m] Day Treatment Center (TTZ), Interdisciplinary Tumor Center Mannheim (ITM) & 3rd Department of Medicine, University Medical Centre Mannheim, Medical Faculty Mannheim of the University of Heidelberg, Theodor-Kutzer-Ufer 1-3, 68167, Mannheim, Germany
[n] 3rd Department of Internal Medicine (Hematology/Medical Oncology), Klinikum Rechts der Isar, Technische Universität München, Ismaninger Str. 22, 81675, Munich, Germany

Received 21 June 2019; received in revised form 25 September 2019; accepted 1 October 2019
Available online 1 November 2019

KEYWORDS
Gastroesophageal adenocarcinoma;
Neoadjuvant chemotherapy;
Neoadjuvant chemoradiotherapy;
Overall survival;
Disease-free survival;
Individual patient data meta-analysis

Abstract ***Introduction:*** Disease-free survival (DFS) is increasingly being used as surrogate end-point for overall survival (OS) in cancer trials. So far, there has been no validation of the surrogacy of DFS for OS for neoadjuvant treatment of gastroesophageal adenocarcinoma.
Methods: The study uses individual patient data (IPD) from eight randomised controlled trials (RCTs) (n = 1126 patients) comparing neoadjuvant therapy followed by surgery with surgery alone for gastroesophageal adenocarcinoma. Correlation between OS time and DFS time was calculated to evaluate individual-level surrogacy. For each trial, survival curves using the Kaplan-Meier method were plotted and hazard ratios (HRs) on the treatment effects were calculated for OS and DFS separately. Those HRs were pooled in a random-effects meta-analysis. Observed HRs were compared with predicted HRs for OS using results from an error-in-variables linear regression model accounting for the uncertainty about the estimated effect. The strength of the association was quantified by the coefficient of determination to assess trial-level surrogacy. The surrogate threshold effect was calculated to determine the minimum treatment effect on DFS necessary to predict a non-zero treatment effect on OS.
Results: A strong correlation between OS time and DFS time was observed, indicating a high individual-level surrogacy. For all RCTs, estimated HRs for OS and DFS were highly similar. In the meta-analysis, the overall HR for OS was virtually identical to that for DFS. The estimated coefficient of determination r^2 for the association between HRs for OS and DFS was 0.912 (95% confidence interval: 0.75–1.0), indicating a very good fit of the regression model and thus a strong trial-level surrogacy between OS and DFS. The surrogate threshold effect based on the regression analysis was 0.79.
Discussion: Based on strong correlations between DFS and OS, as well as a strong correlation of the treatment effects of the two end-points in the error-in-variable regression, DFS seems an appropriate surrogate marker for OS in randomised trials of neoadjuvant chemotherapy or chemoradiotherapy for gastroesophageal adenocarcinoma.

1. Introduction

Most patients with gastroesophageal adenocarcinoma, i.e. adenocarcinoma of the oesophagus, gastroesophageal junction or stomach, present with advanced disease [1,2]. In the absence of distant metastasis, oncological resection is the only potentially curative modality. However, the prognosis after surgery alone in locally advanced disease is poor with 5-year survival rates of only about 20–30%. A considerable proportion of patients who undergo upfront resection will eventually relapse and die as a result of their disease. Multimodal treatment concepts comprising chemotherapy or chemoradiotherapy have shown survival benefits [3–7]. Because of difficulties in administering chemotherapy or radiotherapy soon after surgical procedures, major efforts have been undertaken to explore different neoadjuvant treatment strategies to improve outcomes. Several randomised trials and meta-analyses have demonstrated that neoadjuvant chemotherapy and neoadjuvant chemoradiotherapy followed by surgery are associated with longer survival than surgery alone [3,4]. Therefore, both strategies are now recommended by guidelines for the treatment of non-metastatic gastroesophageal adenocarcinoma [8,9]. Despite this approach, the rate of recurrence remains high, and new neoadjuvant treatment concepts including postoperative chemoradiation after neoadjuvant chemotherapy such as in the CRITICS (ChemoRadiotherapy after Induction chemoTherapy In Cancer of the Stomach) trial [10] and biologicals such as dual HER2 targeting in the INNOVATION (INtegratioN of Trastuzumab, With or Without Pertuzumab, Into periOperatiVe chemotherApy of HER-2 posiTIve

stOmach caNcer) trial [11] or immunotherapy are being explored [12].

The conduction of trials assessing such new treatments is often difficult, as they require a large number of patients to detect the relatively small incremental benefits of a new treatment. Overall survival (OS) has been considered the gold standard end-point for cancer clinical trials. However, OS requires an extended follow-up period and thus trial duration, which leads to higher costs and a long delay until results become available. This has resulted in an increasing number of cancer trials using actuarial end-points, which can be ascertained sooner, such as disease-free survival (DFS) [13]. In non-metastatic gastroesophageal adenocarcinoma, a large proportion of relapses occur before 3 years [3]. However, postrecurrence treatment could dilute or even eliminate an apparent improvement in tumour control achieved by neoadjuvant treatment. Therefore, the validity of DFS as surrogate end-point for OS in gastroesophageal adenocarcinoma remains controversial.

A meta-analysis concluded that DFS is an appropriate end-point for OS in studies of gastric cancer in the adjuvant setting, showing that the effect of treatment on OS is largely predictable from its effect on DFS [14], whereas another analysis found this not to be true in the palliative setting [15]. However, most evidence from large-scale randomised trials in resectable gastroesophageal adenocarcinoma exists for neoadjuvant treatment; which has now, as opposed to adjuvant treatment, become standard of care for non-metastatic disease. Therefore, the present analysis assesses if DFS is a valid surrogate end-point for OS in trials using neoadjuvant chemotherapy or chemoradiotherapy for gastroesophageal adenocarcinoma.

2. Material and methods

2.1. Trial and patient selection

We used individual patient data (IPD) from randomised controlled trials (RCTs) comparing neoadjuvant chemotherapy or chemoradiotherapy (i.e. therapy administered at least partially before surgery) with surgery alone for gastroesophageal adenocarcinoma. All RCTs including patients with resectable, non-metastatic tumours without prior treatment and providing information on both OS and DFS were potentially eligible. There were no exclusion criteria regarding specific treatment regimens. Trials were identified by a systematic literature review covering publications until 2011, the details of which have been previously published [3,16]. All trialists from eligible trials were solicited to provide IPD, and trials were only included in the analyses in case of an affirmative response. Upon data collection, trialists had been asked to provide most recent follow-up data, even if the follow-up period was longer than that reported in respective publications. Some of the IPD had already been used in a previous meta-analysis comparing treatment effects of neoadjuvant chemotherapy with those of surgery alone and in a secondary analysis exploring predictors of postoperative survival [3,17]. From two eligible RCTs [4,18], the final results were only published after completion of that meta-analysis. One of these trials [18] provided IPD for the present analysis. This resulted in IPD from eight trials [18–25], which comprise 1126 patients, entering the analysis (Table 1). All included RCTs had been approved by the respective competent ethical committee.

2.2. Definition of outcomes

OS was defined as time from randomisation to death or to the last documented follow-up. DFS was defined as time from a landmark six months after randomisation to recurrence or death, whichever occurred first, or to the last documented follow-up. The purpose of this landmark was to account for differences in timing between randomisation and surgery between trial arms. Recurrence and death within the first six months were considered events at the landmark.

2.3. Statistical analyses

Characteristics of patients were compared between groups using the chi-square test for discrete variables and the Wilcoxon–Mann–Whitney test for continuous variables. Correlation between OS time and DFS time was assessed by means of the Spearman rank correlation coefficient to assess individual-level surrogacy. OS and DFS were calculated according to the Kaplan-–Meier method separately for patients who received neoadjuvant therapy and patients who underwent surgery alone. This was carried out in the entire study population and for patients from each single RCT. Survival in the single strata was compared using the log-rank test. Hazard ratios (HRs) with 95% confidence intervals were calculated for the comparison of neoadjuvant therapy and surgery alone for each RCT for the treatment effect on both OS and DFS. These HRs were pooled in two separate meta-analyses to provide a combined effect of the estimated HRs. Random-effects models were used for calculation of point estimates and confidence intervals because heterogeneity between the 'true' effects of the different regimens (neoadjuvant chemoradiotherapy or neoadjuvant chemotherapy, different combinations of chemotherapeutical agents, etc.) used in the trials was assumed. In addition, all results were investigated for statistical heterogeneity by I^2 statistics, without using this measure to choose between meta-analytic models.

To compare the observed with the predicted HR for the treatment effect on OS, a linear regression model

Table 1
Randomised controlled trials meeting the inclusion criteria, from which IPD were used in the analysis.

Trial acronym/ first author	Accrual period	Countries	Main inclusion criteria	Chemotherapy/chemoradiotherapy regimen
ACCORD 07 [25]	1995–2003	France (multi-centre)	Adenocarcinoma of the lower third of the oesophagus or GE junction or stomach; UICC stage II or greater; suitable for curative resection; PS 0/1; 18–75 years	2 to 3 cycles (cisplatin 100 mg/m^2 on day 1 or 2; 5-fluorouracil 4000 mg/m^2 cumulative dose over 5 days, then 22 days break) preoperatively; surgery 4–6 weeks after the last chemotherapy dose, 3 to 4 cycles (see above) postoperative 4–6 weeks after surgery for patients who had R0 resection, no progression or major toxicity during preoperative therapy and at least T3 or N+ tumour in histopathology
CALGB 9781 [23]	1997–2000	USA (multi-centre)	Squamous cell or adenocarcinoma of thoracic oesophagus or GE junction, resectable (T1-3, Nx), including regional thoracic lymph node (N1) metastases, supraclavicular lymph node metastasis <1.5 cm, lymph node metastases to levels 15–20 < 1.5 cm; no age limit	1 cycle (cisplatin 200 mg/m^2 cumulative dose on days 1 and 29, 5-fluorouracil 8000 mg/m^2 cumulative dose on days 1–4 and 29 to 32, radiotherapy (1.8 Gy/5 d/wk) begun within 24 h of the chemotherapy administration), continued for 5.5 weeks, final 5.4 Gy given as a boost (total dose 50.4 Gy)
EORTC 40954 [22]	1999–2004	Several European countries, Egypt (multi-centre)	Adenocarcinoma of the stomach or GE junction, cT3/4 Nx M0/ M1 (lymph); PS 0–1; 18–70 years	1 cycle (cisplatin 150 mg/m^2 cumulative dose on days 1, 15 and 29; 5-fluorouracil 12,000 mg/m^2 cumulative dose on days 1, 8, 15, 22, 29 and 36; folinic acid 3000 mg/m^2 cumulative dose on days 1, 8, 15, 22, 29 and 36); restaging, if no progression or toxicity 1 more cycle as described above restarting on day 50; surgery on days 57–63 of the second cycle
FAMTX [20]	1993–1996	The Netherlands (multi-centre)	Adenocarcinoma of the stomach (not cardia); >cT1; resectable with no evidence of distant metastases; PS 0–2; <75 years	2 cycles (methotrexate 1500 mg/m^2 on day 2; 5-fluorouracil 1500 mg/m^2 on day 2; leucovorin 240 or 480 mg, depending on MTX level) cumulative dose on days 3–4; doxorubicin 30 mg/m^2 on day 15; 13 days break); re-staging; in case of response or stable disease another 2 cycles (methotrexate 1500 mg/m^2 on day 2; 5-fluorouracil 1500 mg/m^2 on day 2; leucovorin 240 or 480 mg, depending on MTX level);
FFCD 9901 [18]	2000–2009	France (multi-centre)	Thoracic oesophageal adenocarcinoma or squamous cell carcinoma; suitable for curative resection; cT1/2N0/1 or cT3N0; PS 0–1; <75 years	2 cycles (fluorouracil and cisplatin. FU 800 mg/m^2 per 24 h was administered as a continuous infusion from days 1–4 and 29 to 32. Cisplatin 75 mg/m^2 on day 1 or 2 and day 29 or 30 or, alternatively, 15 mg/m^2 from days 1–5 and 29 to 33), concomitant radiotherapy (45 Gy five fractions per week over 5 weeks).
RTOG 8911 [21]	1990–1995	USA (multi-centre)	Squamous cell or adenocarcinoma of the thoracic oesophagus or GE junction; stage I-III excluding T4 tumours; absence of supraclavicular or distant metastases; fit for surgery; at least 18 years;	3 cycles (cisplatin 100 mg/m^2 on day 1; 5-fluorouracil 1000 mg/m^2 cumulative dose on days 1–5, 23 days break); operation 2–4 weeks after the end of the last cycle; in case of stable or responsive disease upon surgery 2 postoperative cycles (see above, except cisplatin dose reduced to 75 mg/m^2) starting 2–6 weeks after surgery
TROG-AGITG IG9401 [19]	1994–2000	Australia, New Zealand, Singapore (multi-center)	Invasive cancer of the thoracic oesophagus; cT1-3 cN0-1; no involvement of cervical oesophagus or celiac nodes; PS 0 or 1; no age limit	1 cycle (cisplatin 80 mg/m^2 on day 1; 5-fluorouracil 3200 mg/m^2 cumulative dose on days 1–4) with 35 Gy radiotherapy in 15 fractions over 3 weeks, starting concurrently with chemotherapy; surgery 3–6 weeks after completion of radiotherapy; postoperative radiotherapy permitted for patients with residual disease after surgery if indicated clinically for patients assigned to surgery alone
Urba [24]	1989–1994	USA (single-centre)	Squamous cell, adenocarcinoma or mixed adenosquamous carcinoma of the oesophagus or GE junction, limited to the oesophagus and regional lymph nodes (including celiac nodes); Karnofsky index ≥ 60%; ≤75 years	1 cycle (cisplatin 200 mg/m^2 cumulative dose on days 1 through 5 and 17 through 21, 5-fluorouracil 6300 mg/m^2 cumulative on days 1 through 21, vinblastin 8 mg/m^2 on days 1 through 4 and 17 through 20, radiotherapy in fractions of 1.5 Gy twice a day, on days 1 through 5, 8 through 12, and 15 through 19, to a total dose of 45 Gy)

PS: performance status (ECOG/WHO); ECOG: Eastern Cooperative Oncology Group; WHO: World Health Organisation.

accounting for the uncertainty about the estimated effects was used. The treatment effects on DFS were included as predictors in an error-in-variables linear regression model with 95% prediction limits to predict the treatment effects on OS. The strength of the association was quantified by the coefficient of determination r^2 to assess trial-level surrogacy. Considering that the estimated treatment effects from the individual trials on DFS will include a measurement error, we added an additive measurement error in the observed variable [26–29]. The standard error of R^2 was bootstrapped (1000 repetitions), and the 95% confidence interval was then calculated using the bootstrapped standard error and quantiles based on the student's' t-distribution.

To better evaluate the median follow-up requested for future trials using DFS as primary end-point, the correlation coefficient between DFS time and OS time at varying time points (one, two, three, and four years of follow-up) was additionally calculated. To account for the repeated use of data stemming from the same patients for the two end-points, we used the copula approach as described by Rotolo *et al.* as sensitivity analysis [30].

All significance tests were two-sided with $p = 0.05$ as cut-off. IPD were analysed using SAS 9.4 (SAS Institute Inc.). Meta-analyses were conducted using Stata 14.2 (Stata Corp.), and the copula estimation was performed in R, version 3.5.1, using the extension surrosurv [30].

3. Results

Table 1 provides an overview of the characteristics of the eight included RCTs. All but one trial were multicentre trials. Three trials were carried out in the USA, two in France, one in the Netherlands and three in several countries in Europe, North Africa or Australasia. Four RCTs comprised a neoadjuvant chemoradiotherapy scheme and four a neoadjuvant chemotherapy scheme in their experimental arm. All trials used 5-fluorouracil, and seven of eight trials cisplatin as chemotherapeutical backbone. Table 2 shows demographic and clinical characteristics of the 1126 patients included in the analysis, both for the entire study population and separately for patients from the neoadjuvant therapy and surgery-alone arms. Histological and molecular tumour subtypes were not consistently reported in the included trials. Therefore, these data were not available. There were no relevant differences in demographic and preoperative clinical characteristics between the pooled populations from the two study arms. Most patients were male and had a good performance status and a tumour location at the oesophagus or gastroesophageal junction. Postoperatively, patients who had undergone neoadjuvant chemotherapy had significantly less often advanced T and N stages and a significantly higher rate of complete resection.

OS time and DFS time were highly correlated with a Spearman rank correlation coefficient of 0.8943, indicating a good individual-level surrogacy. If follow-up in the investigated trials had been only one, two, three or four years for DFS, the corresponding correlation coefficients would have been 0.68, 0.77, 0.82 and 0.85, respectively. The median follow-up for all patients included in the analysis was 2.10 years (95% confidence interval: 1.92–2.29 years). During follow-up, in the neoadjuvant treatment arms, 389 patients had a recurrence or died without documented prior recurrence, counting as event in the DFS analysis. Three hundred ninety-three patients died, counting as event in the OS analysis. In the surgery-alone arms, 419 patients had an event counting in the DFS analysis and 423 patients in the OS analysis. For 361 patients in the neoadjuvant treatment arms and 398 patients in the surgery-alone arms, recurrence was documented as death of the patient, i.e. either no recurrence was diagnosed before the death of the patient or recurrence was diagnosed at the time of death. In these cases, DFS and OS were the same.

In Fig. 1, curves for OS and DFS, calculated according to the landmark method, are presented stratified by treatment arm. Both OS and DFS are longer in patients who had received neoadjuvant treatment than in those who had undergone surgery alone. OS and DFS curves run largely parallel in patients who had received neoadjuvant treatment as well as in patients who had undergone surgery alone.

Fig. 2 shows the forest plot of HRs for the comparison of treatment effects on OS and DFS between neoadjuvant therapy and surgery alone. There are differences between the absolute values of the HRs from the single RCTs, with the HR indicating a survival benefit for neoadjuvant chemotherapy. For each separate RCT, the point estimate of the HR for the treatment effect on OS and DFS is highly similar. The point estimate and confidence interval of the HR for the treatment effect on OS is virtually identical to those of the HR for the treatment effect on DFS.

Results of the error-in-variable regression are shown in Fig. 3. For one relatively small RCT comparing neoadjuvant chemoradiotherapy (cisplatin/5-fluorouracil with 50.4 Gy) with surgery alone, the observed HR for the treatment effect on OS was lower than expected based on the HR for the treatment effect on DFS [23]. For all other RCTs, the deviation of observed OS from what would be expected based on DFS is small and within the confidence limits. The coefficient of determination r^2 for the association between the HRs for the treatment effects on OS and DFS is 0.912 (95% confidence interval: 0.75–1.0), indicating a very good fit of the regression model and thus a strong trial-level surrogacy between OS and DFS. The

Table 2
Demographic and clinical characteristics of patients included in the analysis, both for the entire study population and separately for patients from the neoadjuvant therapy and surgery-alone arms.

Characteristic	Neoadjuvant Chemotherapy N = 562	Surgery alone N = 564	Total N = 1126	p-value
Trial				
ACCORD 07 [25]	113 (20.1%)	111 (19.7%)	224 (19.9%)	0.998
CALGB 9781 [23]	23 (4.1%)	19 (3.4%)	42 (3.7%)	
EORTC 40954 [22]	69 (12.3%)	71 (12.6%)	140 (12.4%)	
FAMTX [20]	27 (4.8%)	29 (5.1%)	56 (5.0%)	
FFCD 9901 [18]	98 (17.4%)	97 (17.2%)	195 (17.3%)	
RTOG 8911 [21]	115 (20.5%)	121 (21.5%)	236 (21.0%)	
TROG-AGITG IG9401 [19]	80 (14.2%)	78 (13.8%)	158 (14.0%)	
Urba [24]	37 (6.6%)	38 (6.7%)	75 (6.7%)	
Gender				
male	483 (85.9%)	467 (82.8%)	950 (84.4%)	0.147
female	79 (14.1%)	97 (17.2%)	176 (15.6%)	
Age [years]				
N	562	564	1126	0.908
Mean ± SD	59.8 ± 9.3	59.6 ± 9.4	59.7 ± 9.3	
p5, p25, p75, p95	44.0, 53.2, 67.0, 73.2	43.0, 53.3, 67.0, 73.0	44.0, 53.2, 67.0, 73.1	
Median	60.8	61.0	61.0	
Min, Max	23.0, 78.0	26.1, 80.5	23.0, 80.5	
Age				
< 65 years	366 (65.1%)	377 (66.8%)	743 (66.0%)	0.827
65–75 years	184 (32.7%)	176 (31.2%)	360 (32.0%)	
> 75 years	12 (2.1%)	11 (2.0%)	23 (2.0%)	
Tumour location				
Stomach	88 (15.7%)	89 (15.8%)	177 (15.7%)	0.984
GE junction	153 (27.2%)	158 (28.0%)	311 (27.6%)	
Oesophagus	261 (46.4%)	260 (46.1%)	521 (46.3%)	
Oesophagus/GE junction (no further specification)	60 (10.7%)	57 (10.1%)	117 (10.4%)	
Performance status				
0	373 (71.5%)	365 (71.3%)	738 (71.4%)	0.263
1	144 (27.6%)	146 (28.5%)	290 (28.0%)	
2	5 (1.0%)	1 (0.2%)	6 (0.6%)	
missing	40	52	92	
T stage [preoperative, clinical]				
T0	1 (0.5%)	0 (0.0%)	1 (0.2%)	0.867
T1	22 (10.7%)	16 (7.7%)	38 (9.2%)	
T2	57 (27.7%)	56 (27.0%)	113 (27.4%)	
T3	121 (58.7%)	130 (62.8%)	251 (60.8%)	
T4	5 (2.4%)	5 (2.4%)	10 (2.4%)	
missing	357	361	718	
N stage [preoperative, clinical]				
N0	61 (80,8%)	47 (63.5%)	108 (70.1%)	0.178
N1	18 (22.5%)	27 (36.5%)	45 (29.2%)	
N2	1 (1.3%)	0 (0.0%)	1 (0.6%)	
missing	482	490	972	
T stage [postoperative, histopathological]				
T0	53 (13.2%)	2 (0.5%)	55 (6.7%)	<0.001
T1	63 (15.7%)	64 (15.2%)	127 (15.5%)	
T2	112 (27.9%)	106 (25.2%)	218 (30.2%)	
T3	156 (38.9%)	207 (49.2%)	363 (50.3%)	
T4	17 (4.2%)	42 (10.0%)	59 (8.2%)	
missing	161	143	304	
N stage [postoperative, histopathological]				
N0	181 (45.3%)	110 (26.4%)	291 (35.6%)	<0.001
N1	171 (42.8%)	210 (50.4%)	381 (46.6%)	
N2	35 (8.8%)	59 (14.1%)	94 (11.5%)	
N3	13 (3.3%)	38 (9.1%)	51 (6.2%)	
missing	162	147	309	
Margin status				
R0	395 (91.2%)	374 (82.3%)	769 (86.7%)	0.001
R1	18 (4.2%)	35 (7.4%)	53 (6.0%)	
R2	20 (4.6%)	45 (9.5%)	65 (7.3%)	
missing	129	110	239	

SD: standard deviation.

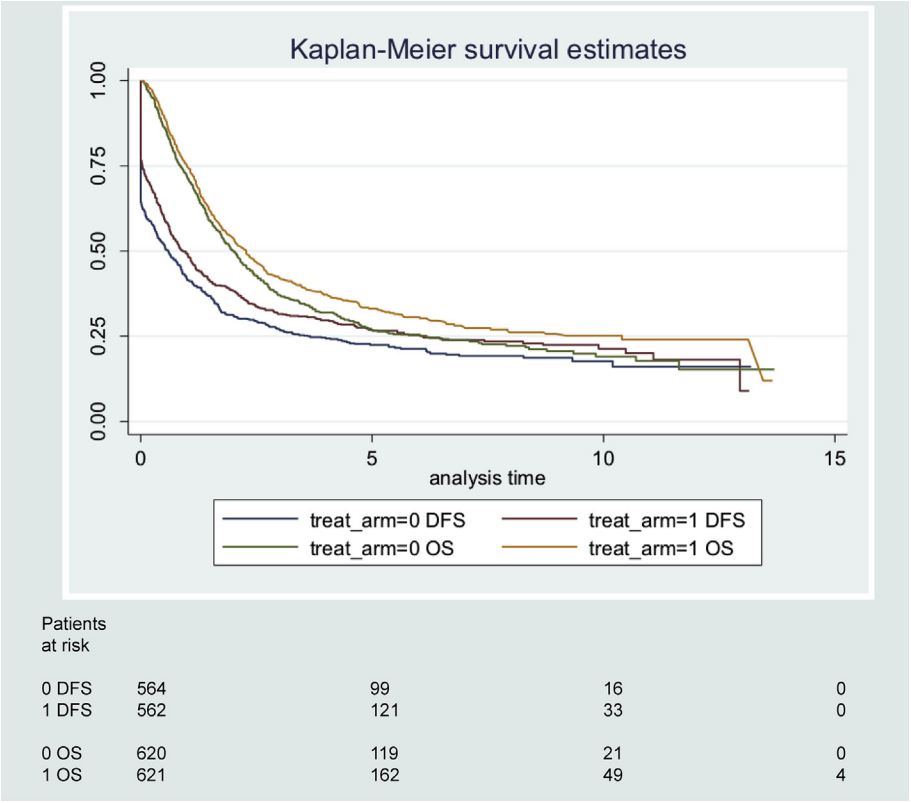

Fig. 1. Time-to-event curves for OS and DFS, calculated according to the landmark method, stratified by treatment arm. Treat_arm = 0 OS: overall survival in the upfront surgery arms, treat_arm = 1 OS: overall survival in the neoadjuvant therapy arms, treat_arm = 0 DFS: disease-free survival in the upfront surgery arms, treat_arm = 1 DFS: disease-free survival in the neoadjuvant therapy arms. OS: overall survival; DFS: disease-free survival.

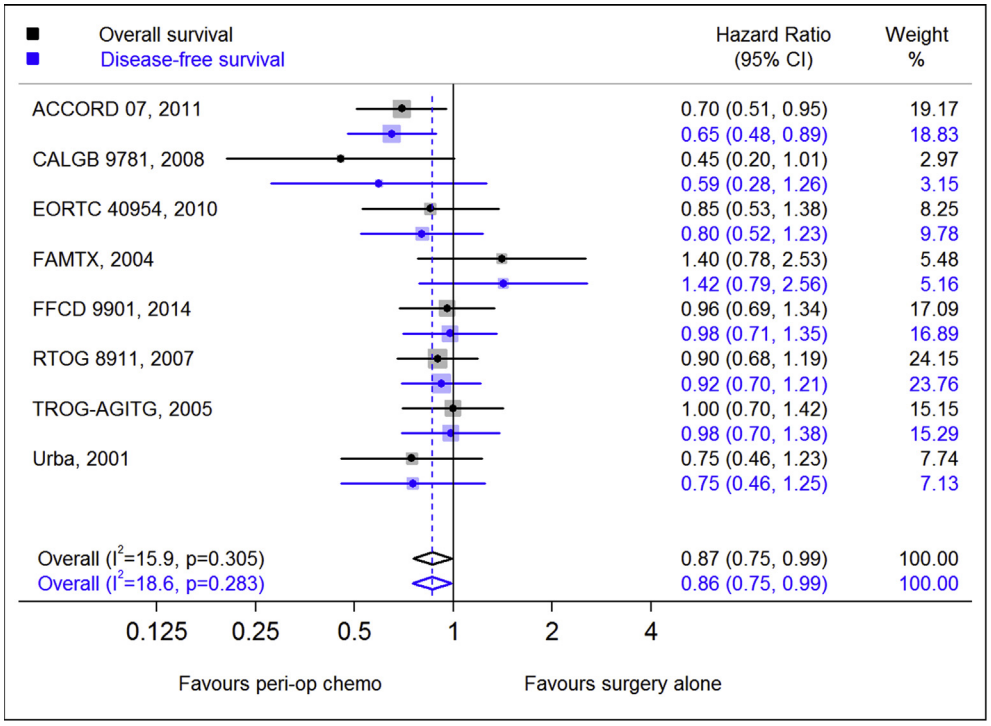

Fig. 2. Forest plot of hazard ratios for the comparison of OS and DFS between neoadjuvant therapy and surgery alone. CI: confidence interval; OS: overall survival; DFS: disease-free survival.

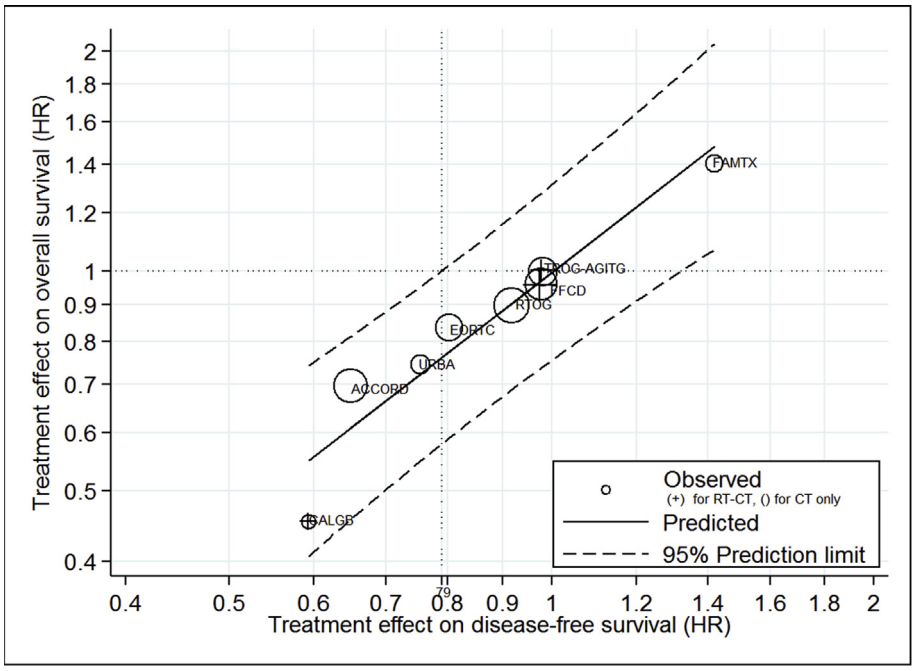

Fig. 3. Error-in-variable regression, observed and predicted HRs for OS and DFS with 95% prediction limits. RT-CT: neoadjuvant chemoradiotherapy, CT: neoadjuvant chemotherapy; OS: overall survival; DFS: disease-free survival; HR: hazard ratio.

surrogate threshold effect based on the regression analysis was 0.79. This means that a future trial yielding a HR for the treatment effect on DFS below 0.79 could be expected with a 95% probability to yield a HR for the treatment effect on OS below one.

In addition to the regression analyses, copula estimation was performed to account for correlation between the treatment effects of the two outcomes. Results using an unadjusted Clayton copula are very similar to those of the regression analysis. The coefficient of determination for the copula analysis is 0.95, while the surrogate threshold effect is 0.74. As convergence could not be achieved, adjustment for the second-step linear regression for measurement error in the copula model was not performed.

4. Discussion

The aim of this IPD analysis was to assess individual- and trial-level surrogacy between OS and DFS, or in other words, how predictable OS was by DFS in randomised trials comparing neoadjuvant treatment with surgery alone for gastroesophageal adenocarcinoma. In case of a strong and reliable prediction, DFS could be used as a valid surrogate end-point, shortening overall trial duration and providing trial results faster.

The results show a considerable correlation between the two outcomes in patients from the included trials, both between the two outcomes themselves (individual-level surrogacy) and between the treatment effects on the outcomes estimated in the individual trials (trial-level surrogacy). This observation can partially be explained by the fact that 57% of patients died without prior diagnosis of recurrence, which led to their DFS being identical to their OS. The standard definition of DFS used in oncological trials comprises that deaths without documented prior recurrence are counted as events in DFS analyses [31]. Thus, this finding is externally valid with regard to other trials. Most patients with disease recurrence, however, died several months or few years after diagnosis of the recurrence. The length of the time interval between recurrence and death is potentially influenced by chemotherapeutic, radiotherapeutic or even surgical treatment. None of the trials provided information on an individual patient level if and what treatment patients received. Therefore, one can only speculate about its possible effects. In general, it might be assumed that patients who underwent surgery alone and are thus chemotherapy-naïve receive more dose-intense chemotherapy than patients who had already undergone neoadjuvant chemotherapy. Similarly, radiotherapy for locoregional recurrence can usually only be administered in patients who had not been treated with neoadjuvant irradiation. However, given the biological complexity of the disease, it cannot be readily concluded that the time between recurrence and death is indeed longer in patients without prior neoadjuvant therapy.

The correct determination of DFS strongly depends on follow-up intervals and the specific kind of clinical, radiological and histopathological examinations conducted to detect disease recurrence. These inevitably vary across the RCTs included in our analyses because they were carried out in different settings and during

different time periods. However, as visual inspection resulted in no violation of the proportional hazards assumption, and HRs can therefore be assumed to be time-independent, length, intensity and frequency of follow-up will not influence the estimation. OS, on the one hand, is a very stable indicator because ascertainment of death during follow-up is supposedly very accurate. The consistency of the correlation between DFS and OS across trials does not suggest large differences in DFS determination across the different trials. The choice of the landmark time at six months after randomisation for DFS analyses is somewhat arbitrary. Early therapy-related and postoperative deaths as well as deaths during the early phase of postoperative continuation of chemotherapy might not be mirrored exactly by this approach. On the other hand, these deaths are all accounted for in the OS analysis. There is no commonly agreed landmark time for such analyses, but six months have been used in an RCT on neoadjuvant chemotherapy for oesophageal squamous cell cancer [32] and in our previous meta-analyses [3,17], which led us to choose the same value for the present analysis.

In the meta-analysis, the overall HRs for the treatment effects on OS and DFS are virtually identical. Similarly, the coefficient of determination in the error-in-variable regression is close to one. This indicates a very good model fit which reflects a strong correlation between treatment effects on OS and DFS (trial-level surrogacy). Furthermore, the correlation between the two outcomes themselves was also high, indicating a good individual-level surrogacy. These results are consistent across all different included trials, regardless of the specificities of the applied neoadjuvant chemotherapy and regardless of patients receiving combined chemoradiotherapy or merely chemotherapy. Only one small trial using radiotherapy along with cisplatin/5-fluorouracil doublet chemotherapy constituted an outlier, as the observed HR for the treatment effect on DFS was lower than what was expected based on OS in the trial [23].

The results indicate that most variations in OS can be explained by the effect of the respective neoadjuvant treatment on DFS. This speaks in favour of DFS serving as an appropriate surrogate marker for OS in trials evaluating neoadjuvant chemotherapy or chemoradiotherapy in patients with gastroesophageal adenocarcinoma. As a limitation, none of the included trials used targeted therapy with biologicals or monoclonal antibodies, or immunotherapy, and therefore, deductions regarding these more recent treatment concepts [11,12] are difficult to make. Moreover, histological and molecular tumour subtypes were not consistently reported in the included trials and therefore not available for analysis. Response to therapy depends on these characteristics, and although there is no direct evidence to that regard, the association between DFS and OS might be influenced by the histological or molecular subtype of the tumour.

The calculated surrogate threshold effect of 0.79 means that in future trials, a treatment yielding a reduction in the hazards of disease recurrence of at least 21% can be assumed to have a beneficial effect also on overall survival. Comparable analyses have found a surrogate threshold effect of 0.92 for adjuvant chemotherapy in gastric cancer [14] and of 0.56 for chemotherapy in metastatic or irresectable gastric cancer [15]. There is no clearly established limit above which a surrogate threshold effect would qualify an outcome as an appropriate surrogate outcome for another one. However, most RCTs in oncology are powered to detect effects in the magnitude of a 20%–30% reduction in the hazards of relapse or death. Moreover, based on an individual patient-level analysis of multiple randomised trials by Shi *et al.* [33], 0.8 is often regarded as a meaningful boundary. Therefore, with a surrogate threshold effect of 0.79 in the present analysis, DFS can be regarded as a reasonably appropriate surrogate endpoint for OS. It must be noted, however, that methodologically, the surrogate threshold effect is computed for trials of infinite size. Given that the included RCTs were all of medium size, the surrogate threshold effect might be overestimated.

In summary, based on a strong correlation between DFS and OS on both the individual patient and trial level, as well as on the finding that the vast majority of variation in OS can be explained by variation in DFS, DFS seems to be an appropriate surrogate marker for OS in randomised trials of neoadjuvant chemotherapy or chemoradiotherapy for gastroesophageal adenocarcinoma. However, as novel treatment concepts with substances other than cytotoxic compounds keep evolving, this finding requires continued validation.

Conflict of interest statement

None declared.

Acknowledgements

Parts of this work were supported by the German Federal Ministry of Education and Research (grant number 01 KG 0807) and intramural research funds of the University of Heidelberg (Foundation for Cancer Research). The funding sources had no role in the study design; collection, analysis and interpretation of data; in the writing of the report and in the decision to submit the article for publication. This study was conducted in memory of Prof. Christophe Mariette who initiated the study FFCD9901. The authors thank Emilie Barbier, FFCD Statistician, for her support.

References

[1] Rubenstein JH, Shaheen NJ. Epidemiology, diagnosis, and management of esophageal adenocarcinoma. Gastroenterology 2015;149(2). 302-317.e1.

[2] Jim MA, Pinheiro PS, Carreira H, Espey DK, Wiggins CL, Weir HK. Stomach cancer survival in the United States by race and stage (2001-2009): findings from the CONCORD-2 study. Cancer 2017;123(Suppl 24):4994–5013.

[3] Ronellenfitsch U, Schwarzbach M, Hofheinz R, Kienle P, Kieser M, Slanger TE, et al. Preoperative chemo(radio)therapy versus primary surgery for gastroesophageal adenocarcinoma: systematic review with meta-analysis combining individual patient and aggregate data. Eur J Cancer 2013;49(15):3149–58 (Oxford, England : 1990).

[4] van Hagen P, Hulshof MC, van Lanschot JJ, Steyerberg EW, van Berge Henegouwen MI, Wijnhoven BP, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 2012;366(22):2074–84.

[5] Schirren R, Reim D, Novotny AR. Adjuvant and/or neoadjuvant therapy for gastric cancer? A perspective review. Ther Adv Med Oncol 2015;7(1):39–48.

[6] Egyud MR, Tseng JF, Suzuki K. Multidisciplinary therapy of esophageal cancer. Surg Clin N Am 2019;99(3):419–37.

[7] Al-Batran SE, Homann N, Pauligk C, Goetze TO, Meiler J, Kasper S, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. Lancet (London, England) 2019;393(10184):1948–57.

[8] Lordick F, Mariette C, Haustermans K, Obermannova R, Arnold D. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol : Off J Eur Soc Med Oncol 2016;27(suppl 5):v50–7.

[9] Smyth EC, Verheij M, Allum W, Cunningham D, Cervantes A, Arnold D. Gastric cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol : Off J Eur Soc Med Oncol 2016;27(suppl 5):v38–49.

[10] Cats A, Jansen EPM, van Grieken NCT, Sikorska K, Lind P, Nordsmark M, et al. Chemotherapy versus chemoradiotherapy after surgery and preoperative chemotherapy for resectable gastric cancer (CRITICS): an international, open-label, randomised phase 3 trial. Lancet Oncol 2018;19(5):616–28.

[11] Wagner AD, Grabsch HI, Mauer M, Marreaud S, Caballero C, Thuss-Patience P, et al. EORTC-1203-GITCG - the "INNOVATION"-trial: effect of chemotherapy alone versus chemotherapy plus trastuzumab, versus chemotherapy plus trastuzumab plus pertuzumab, in the perioperative treatment of HER2 positive, gastric and gastroesophageal junction adenocarcinoma on pathologic response rate: a randomized phase II-intergroup trial of the EORTC-Gastrointestinal Tract Cancer Group, Korean Cancer Study Group and Dutch Upper GI-Cancer group. BMC Canc 2019;19(1):494.

[12] Maron SB, Catenacci DVT. Update on gastroesophageal adenocarcinoma targeted therapies. Hematol Oncol Clin N Am 2017;31(3):511–27.

[13] Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, Kramar A. Survival end point reporting in randomized cancer clinical trials: a review of major journals. J Clin Oncol : Off J Am Soc Clin Oncol 2008;26(22):3721–6.

[14] Oba K, Paoletti X, Alberts S, Bang YJ, Benedetti J, Bleiberg H, et al. Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. J Natl Cancer Inst 2013;105(21):1600–7.

[15] Paoletti X, Oba K, Bang YJ, Bleiberg H, Boku N, Bouche O, et al. Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. J Natl Cancer Inst 2013;105(21):1667–70.

[16] Ronellenfitsch U, Schwarzbach M, Hofheinz R, Kienle P, Kieser M, Slanger TE, et al. Perioperative chemo(radio)therapy versus primary surgery for resectable adenocarcinoma of the stomach, gastroesophageal junction, and lower esophagus. Cochrane Database Syst Rev 2013;(5):Cd008107.

[17] Ronellenfitsch U, Schwarzbach M, Hofheinz R, Kienle P, Nowak K, Kieser M, et al. Predictors of overall and recurrence-free survival after neoadjuvant chemotherapy for gastroesophageal adenocarcinoma: pooled analysis of individual patient data (IPD) from randomized controlled trials (RCTs). Eur J Surg Oncol : J Eur Soci Surg Oncol Br Assoc Surg Oncol 2017;43(8):1550–8.

[18] Mariette C, Dahan L, Mornex F, Maillard E, Thomas PA, Meunier B, et al. Surgery alone versus chemoradiotherapy followed by surgery for stage I and II esophageal cancer: final analysis of randomized controlled phase III trial FFCD 9901. J Clin Oncol: Off J Am Soc Clin Oncol 2014;32(23):2416–22.

[19] Burmeister BH, Smithers BM, Gebski V, Fitzgerald L, Simes RJ, Devitt P, et al. Surgery alone versus chemoradiotherapy followed by surgery for resectable cancer of the oesophagus: a randomised controlled phase III trial. Lancet Oncol 2005;6(9):659–68.

[20] Hartgrink HH, van de Velde CJ, Putter H, Songun I, Tesselaar ME, Kranenbarg EK, et al. Neo-adjuvant chemotherapy for operable gastric cancer: long term results of the Dutch randomised FAMTX trial. Eur J Surg Oncol : J Eur Soci Surg Oncol Br Assoc Surg Oncol 2004;30(6):643–9.

[21] Kelsen DP, Winter KA, Gunderson LL, Mortimer J, Estes NC, Haller DG, et al. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. J Clin Oncol : Off J Am Soc Clin Oncol 2007;25(24):3719–25.

[22] Schuhmacher C, Gretschel S, Lordick F, Reichardt P, Hohenberger W, Eisenberger CF, et al. Neoadjuvant chemotherapy compared with surgery alone for locally advanced cancer of the stomach and cardia: european Organisation for Research and Treatment of Cancer randomized trial 40954. J Clin Oncol : Off J Am Soc Clin Oncol 2010;28(35):5210–8.

[23] Tepper J, Krasna MJ, Niedzwiecki D, Hollis D, Reed CE, Goldberg R, et al. Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. J Clin Oncol : Off J Am Soc Clin Oncol 2008;26(7):1086–92.

[24] Urba SG, Orringer MB, Turrisi A, Iannettoni M, Forastiere A, Strawderman M. Randomized trial of preoperative chemoradiation versus surgery alone in patients with locoregional esophageal carcinoma. J Clin Oncol : Off J Am Soc Clin Oncol 2001;19(2):305–13.

[25] Ychou M, Boige V, Pignon JP, Conroy T, Bouche O, Lebreton G, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol : Off J Am Soc Clin Oncol 2011;29(13):1715–21.

[26] Draper N, Smith H. Applied regression analysis. 3rd ed. New York, NY: John Wiley & Sons; 1998.

[27] Kmenta J. Elements of econometrics. 2nd ed. Ann Arbor, MI: University of Michigan Press; 1997.

[28] Treiman D. Quantitative data analysis: doing social research to test ideas. San Francisco, CA: Jossey-Bass; 2008.

[29] Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. Stat Methods Med Res 2008;17(5):467–75.

[30] Rotolo F, Paoletti X, Michiels S. surrosurv: an R package for the evaluation of failure time surrogate endpoints in individual patient data meta-analyses of randomized clinical trials. Comput Methods Progr Biomed 2018;155:189–98.

[31] Panageas KS, Knezevic A. Selection of endpoints. In: Halabi S, Michiels S, editors. Textbook of clinical trials in oncology: a statistical perspective. Boca Raton, FL: CRC Press; 2019.

[32] Boonstra JJ, Kok TC, Wijnhoven BP, van Heijl M, van Berge Henegouwen MI, Ten Kate FJ, et al. Chemotherapy followed by surgery versus surgery alone in patients with resectable oesophageal squamous cell carcinoma: long-term results of a randomized controlled trial. BMC Canc 2011;11:181.

[33] Shi Q, Schmitz N, Ou FS, Dixon JG, Cunningham D, Pfreundschuh M, et al. Progression-free survival as a surrogate end point for overall survival in first-line diffuse large B-cell lymphoma: an individual patient-level analysis of multiple randomized trials (SEAL). J Clin Oncol : Off J Am Soc Clin Oncol 2018;36(25):2593–602.