DOI: 10.1093/jnci/djt337 JNCI Journal of the National Cancer Institute Advance Access published October 9, 2013

REVIEW

Disease-Free Survival as a Surrogate for Overall Survival in Adjuvant Trials of Gastric Cancer: A Meta-Analysis

Koji Oba, Xavier Paoletti, Steven Alberts, Yung-Jue Bang, Jacqueline Benedetti, Harry Bleiberg, Paul Catalano, Florian Lordick, Stefan Michiels, Satoshi Morita, Yasuo Ohashi, Jean-pierre Pignon, Philippe Rougier, Mitsuru Sasako, Junichi Sakamoto, Daniel Sargent, Kohei Shitara, Eric Van Cutsem, Marc Buyse, Tomasz Burzykowski; on behalf of the GASTRIC group

Manuscript received February 12, 2013; revised July 25, 2013; accepted July 25, 2013.

Correspondence to: Koji Oba, PhD, Translational Research and Clinical Trial Center, Hokkaido University Hospital, Kita 14, Nishi 5, Kita-ku, Sapporo, Hokkaido 0608648, Japan (e-mail: k.oba@huhp.hokudai.ac.jp).

Background In investigations of the effectiveness of surgery and adjuvant chemotherapy for gastric cancers, overall survival (OS) is considered the gold standard endpoint. However, the disadvantage of using OS as the endpoint is that it requires an extended follow-up period. We sought to investigate whether disease-free survival (DFS) is a valid surrogate for OS in trials of adjuvant chemotherapy for gastric cancer.

Methods The GASTRIC group initiated a meta-analysis of individual patient data collected in randomized clinical trials comparing adjuvant chemotherapy vs surgery alone for patients with curatively resected gastric cancer. Surrogacy of DFS was assessed through the correlation between the endpoints as well as through the correlation between the treatment effects on the endpoints. External validation of the prediction based on DFS was also evaluated.

Results Individual patient data from 14 randomized clinical trials that included a total of 3288 patients were analyzed. The rank correlation coefficient between DFS and OS was 0.974 (95% confidence interval [CI] = 0.971 to 0.976). The coefficient of determination between the treatment effects on DFS and on OS was as high as 0.964 (95% CI = 0.926 to 1.000), and the surrogate threshold effect based on adjusted regression analysis was 0.92. In external validation, the six hazard ratios for OS predicted according to DFS were in very good agreement with those actually observed for OS.

Conclusions DFS is an acceptable surrogate for OS in trials of cytotoxic agents for gastric cancer in the adjuvant setting.

J Natl Cancer Inst

Gastric cancer is the fourth most common malignancy in the world, affecting 989 000 patients in 2008 (7.8% of all cancers) (1). The most effective treatment for localized disease is surgery, but even after curative resection, recurrence is noted in more than half the cases of advanced-stage disease. This poor outcome has prompted major efforts to explore different adjuvant therapies. However, over the last three decades, despite some successful large-scale trials (2–5), only modest improvement has been achieved in survival. Our group recently reported the results of a meta-analysis of individual data that showed a lower risk of death with postoperative adjuvant chemotherapy than with surgery alone (overall hazard ratio [HR] =0.82; $P < .0001$) (6). However, the efficacy of adjuvant chemotherapy is still far from satisfactory, and further investigation into more effective treatments for patients with resectable gastric cancer is warranted.

Historically, the 5-year overall survival (OS) rate has typically been the most quoted metric for judging the success of a particular treatment. This endpoint has the advantage of being simple to measure, easy to interpret, and clinically meaningful. However, the main disadvantages of this endpoint are that it requires an extended follow-up period and its measurement is potentially diluted by nonmalignant causes of death and therapies for recurrent/advanced disease.

A reasonable candidate for a surrogate of OS in the adjuvant setting is disease-free survival (DFS), which is defined here as the time to cancer recurrence, second cancer, or death from any cause. Recent meta-analyses have been used to validate DFS as a surrogate for OS in other tumor types (7,8). If DFS could replace OS in the assessment of the efficacy of new treatments in clinical trials testing adjuvant treatment for patients with curatively resected gastric cancer, the trial duration and costs would be reduced. We performed a comprehensive meta-analysis of data from 3838 individual patients randomized in 17 trials on curatively resected gastric cancer; documented DFS values, which were available for 3371 of the patients from 14 trials, were used to evaluate DFS as a surrogate endpoint for OS.

Methods

Study Selection

Our analyses were based on a meta-analysis of individual patient data (IPD) described in detail elsewhere (6). IPD from all randomized trials comparing adjuvant chemotherapy with surgery alone for resectable gastric cancers were sought electronically from MEDLINE, the Cochrane Central Register of Controlled Trials,

and the National Institutes of Health trial registry (ClinicalTrials.gov). Trials were eligible if they were randomized, closed to patient accrual before 2004, and compared any adjuvant therapy after curative resection with surgery alone.

Data and Outcomes

The following data were requested for all individual patients included in all the trials: center, randomization date, treatment allocated by randomization, date of last follow-up or death, survival status, cause of death (if applicable), relapse status, and type and date of relapse if any. OS was defined as the time from randomization to all-cause death or the date of the last follow-up used for censoring. DFS was defined as the time to relapse, second cancer, or all-cause death, whichever came first. Detailed information on the type of relapse was not always available. All data were centrally reanalyzed and checked for inconsistencies. In particular, diagnostic tools for randomization quality were systematically applied (6).

Statistical Methods

Forest plots were used to display the hazard ratios (HRs) for overall and individual trials, which were then used for the evaluation of surrogacy of DFS for OS (labeled "training trials" in Figure 1) and for external validation trials (labeled "validation literature data" and "validation trials IPD" in Figure 1). The hazard ratios compared the hazard of an event in patients treated with adjuvant chemotherapy with the hazard in patients treated with surgery alone.

We used the Spearman rank correlation coefficient between DFS and OS to assess surrogacy at the individual level and the coefficient of determination between the natural logarithm of the hazard ratios for DFS and OS to assess surrogacy at the trial level (7,9). At the individual level, the association between the distribution of the true endpoint (OS) and the surrogate (DFS) was evaluated using a bivariable model based on the Plackett copula combined with trial-specific Weibull models for DFS and OS (10,11). The association between the estimates of treatment effects obtained using the bivariable model was used to assess surrogacy at the trial level. A good surrogate was considered to provide a reliable prediction of the treatment effect on the true endpoint (eg, the hazard ratio for OS) from the treatment effect on the surrogate (eg, the hazard ratio for DFS). It should be noted that estimates of the hazard ratios based on the bivariable model might differ from the crude estimates shown in the forest plot.

To quantify the association between the natural logarithm of the hazard ratios for OS and DFS, we used a linear regression model that accounted for the uncertainty about the estimated effects by using an error-in-variables linear regression model. The strength of the association was assessed by using the coefficient of determination R^2 (or explained variation). This approach has been previously used for resectable colorectal and metastatic breast cancers (7,12).

Sensitivity Analyses

To assess the typical trial conditions, we performed a sensitivity analysis by studying the association between the treatment effects on OS at 5 years and DFS at different time points (2 years, 3 years, and 4 years), while censoring all events occurring after these time points. Because only durations (OS and DFS) and not dates were provided for two studies, the same individual follow-up was used for all patients, irrespective of their actual accrual date. In this analysis, the number of observed events is considerably lower than that in the analysis of patients followed-up to a common administrative censoring date. In the latter case, analysis takes place 2, 3, or 4 years after the accrual of the last patients. Therefore, the first accrued patient may have much longer follow-up.

External Validation

To assess the external validity of our results, we used 4 trials for which we did not receive IPD from the principal investigators [no reply or refusal to share the data (13, 14) or data lost (15)] and the large-scale CLASSIC trial for which only interim analysis was available at the time of the surrogate analysis (2). We extracted DFS and OS from the summary statistics published for these trials (16). We also used the IPD from a large trial investigating the effect of adjuvant treatment with S1 (TS-1, Taiho Pharmaceutical Company of Bejing Ltd, Bejing, China) vs surgery alone (5) and from a trial studying the benefit of postoperative chemoradiation vs surgery (4).

Surrogate Threshold Effect

On the basis of a linear regression model adjusted for estimation error in observed treatment effects, we calculated the surrogate threshold effect (STE), defined as the minimum treatment effect on DFS necessary to predict a nonzero effect on OS (17). A future trial would require the upper limit of the confidence interval for the estimated hazard ratio for DFS to fall below the STE to predict a nonzero effect on OS.

All analyses were performed on an intention-to-treat basis. Confidence intervals (CI) were calculated for a two-sided probability coverage of 95%. All analyses were performed using SAS software v9.3 (SAS Institute Inc., Cary, NC) except for the graphical displays (double forest plots were plotted using a set of R functions developed at the International Drug Development Institute [Louvain-la-Neuve, Belgium], whereas other figures were prepared using STATA v12 [StataCorp LP, College Station, TX]).

Results

Data were obtained on 3371 patients from the 14 eligible randomized trials with documented OS and DFS (18–30). Nonmissing data on both endpoints were available for 3288 patients, of whom 1763 had events related to DFS and 1705 died during follow-up. Detailed information about treatment regimens and median follow-up studies is provided in Supplementary Table 1 (available online). Figure 1 shows a forest plot of the treatment effects on OS and DFS for all trials. Figure 2 shows the overall Kaplan–Meier curves for DFS and OS. Overall and at the trial level, the effect of any adjuvant chemotherapy on DFS appeared close to the effect on OS ($HR_{OS} = 0.86$; $HR_{DFS} = 0.82$).

Individual- and Trial-Level Association

The individual-level association, as measured by the Spearman rank correlation coefficient, was as high as 0.974 (95% CI = 0.971 to 0.976), indicating a very strong correlation between DFS and OS for a given patient.

A high correlation was noted between log HR_{OS} and log HR_{DFS} (Figure 3). The coefficient of determination, R^2, for the estimated treatment effects was 0.964 (95% CI = 0.926 to 1.000) and 1.000

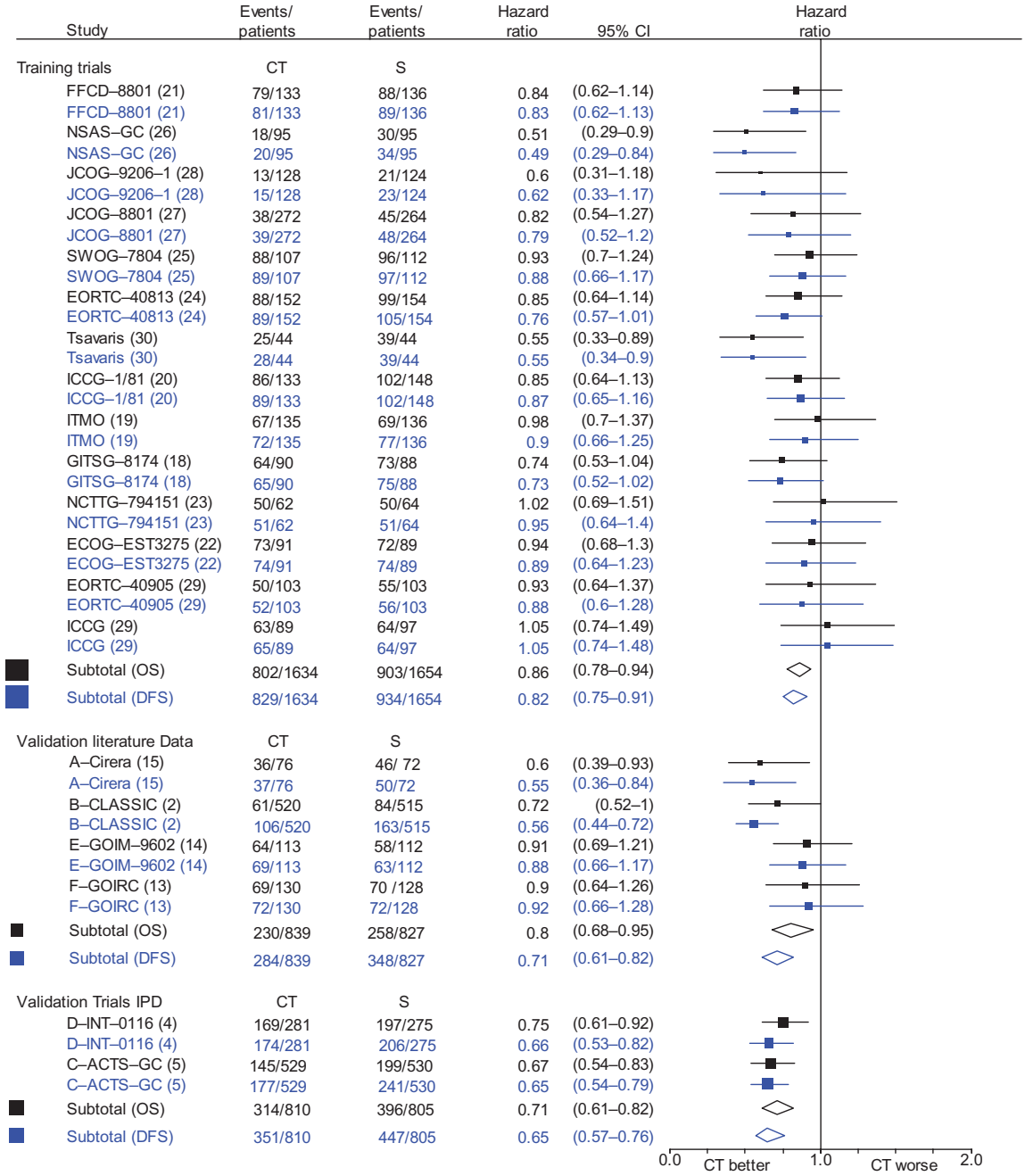

Figure 1. Forest plot of treatment effects (hazard ratios) on disease-free survival (DFS) and on overall survival (OS). The **first row** for each trial shows the result for OS, and the **second row** shows the result of DFS. The **squares** and **diamonds** represent the point estimates and pooled estimates, respectively. **Sizes of the symbols** represent the number of events. The **horizontal error bars** show the 95% confidence interval (CI) of each hazard ratio. CT = adjuvant chemotherapy; IPD = individual patient data; S = surgery alone.

(95% CI = 0.999 to 1.000) before and after adjusting for the estimation error, respectively. Notably, however, because the estimated R^2 value was very close to the upper limit of 1, the obtained numerical results need to be interpreted with caution, as they can be easily influenced by numerical errors.

The linear regression model adjusted for estimation errors was as follows:

$$\ln(\mathrm{HR}_{\mathrm{OS}}) = 0.047 + 1.239 \times \ln(\mathrm{HR}_{\mathrm{DFS}}).$$

In the equation, $\ln(\mathrm{HR}_{\mathrm{OS}})$ and $\ln(\mathrm{HR}_{\mathrm{DFS}})$ denote the natural log transformation of the hazard ratio for each endpoint. Standard errors were 0.023 and 0.151 for the intercept and slope, respectively. This is shown as a straight line on Figure 3, where the x-axis represents the treatment effect on DFS and the y-axis represents the treatment effect on OS. Each trial is represented by a bubble of a size proportional to the trial sample size. The 95% prediction limits indicate the range of effect on OS that can be expected for a given effect on DFS.

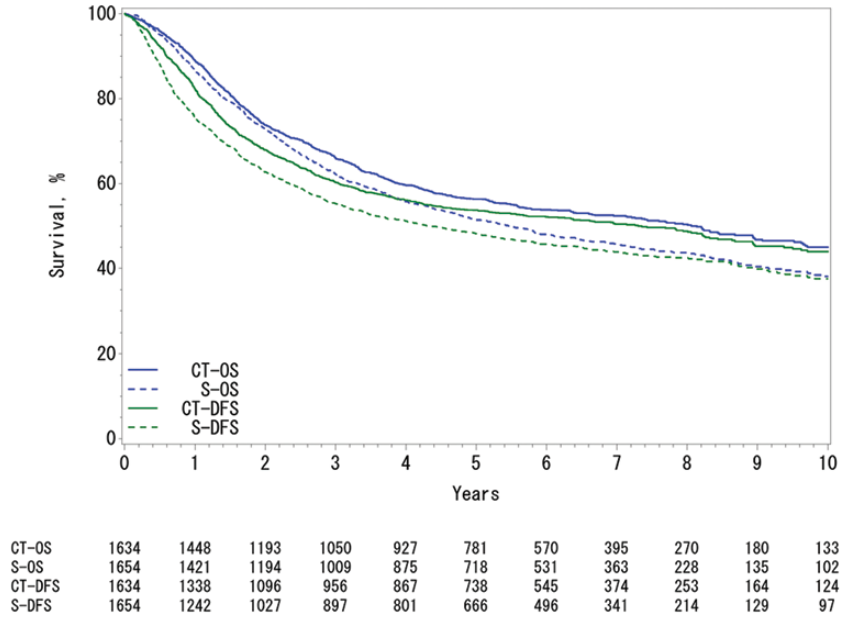

Figure 2. Disease–free survival (DFS) and overall survival (OS) Kaplan–Meier survival curves truncated at 10 years. The number of patients at risk in each group is given below the graph. CT = adjuvant chemotherapy; S = surgery alone.

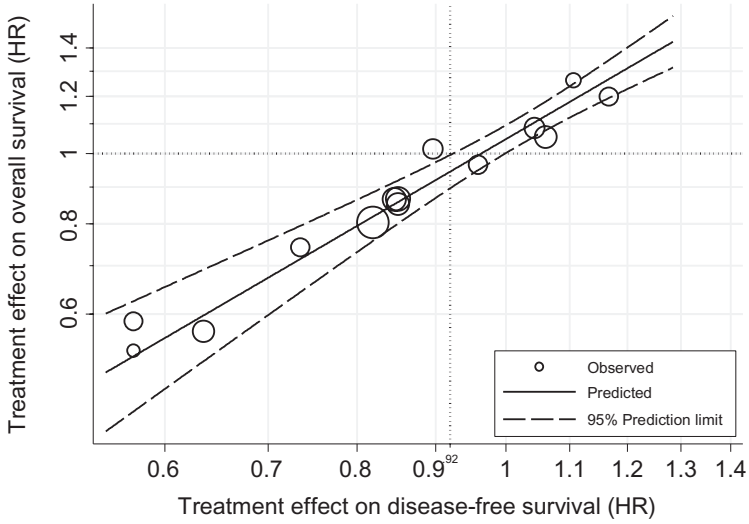

Figure 3. Trial-level association between treatment effects. Log scale was used for the x-axis and y-axis. The **horizontal line** (**dots**) corresponds to the hazard ratio (HR) on overall survival (OS) of 1—that is, the absence of effect on the OS. The **vertical line** (**dots**) crosses the upper boundary of the 95% prediction limit at the point hazard ratio on OS equal to 1. This indicates the surrogate threshold effect.

Considering the high correlation at both the individual and the trial levels, we also computed the STE based on the adjusted regression model. The STE is defined as the intersection of the upper prediction limit, with the horizontal line representing a hazard ratio of 1 for OS (null hypothesis). The STE was equal to 0.92; hence, in a future trial using similar treatment modalities, as in our set of trials, a hazard ratio for DFS less than 0.92 would predict with 95% probability a hazard ratio for OS less than 1.

Sensitivity Analyses

Table 1 shows the association between OS and DFS measured at different time points, ranging from 2 to 4 years after randomization. For this analysis, we report the number of the events available for both OS at 5 years and DFS at 2, 3, and 4 years. Note that at 2 and 3 years, the number of events for DFS was actually lower than that for OS at 5 years, which resulted in wider confidence intervals. Therefore, for this sensitivity analysis, we present values unadjusted for the estimation error.

External Validation

Table 2 and Figure 4 show the results of the external validation using summary data (2,13–15) and IPD (4,5) for six trials. The table displays the observed hazard ratios for OS and DFS with 95% confidence intervals and the hazard ratio for OS predicted from the model of Figure 3 with the 95% predictive prediction intervals. Notably, the 95% confidence interval quantifies the uncertainty of the estimates of

Table 1. Correlation between survival endpoints and surrogacy measures quantification based on the individual patient data from 14 randomized controlled trials*

Summary measures	2-year DFS/5-year OS	3-year DFS/5-year OS	4-year DFS/5-year OS	All
Events	1135/1489	1379/1489	1511/1489	1763/1705
Rho† (95% CI)	0.949 (0.943 to 0.955)	0.953 (0.948 to 0.958)	0.957 (0.952 to 0.961)	0.974 (0.971 to 0.976)
Unadjusted R^2 (95% CI)‡	0.776 (0.569 to 0.983)	0.866 (0.736 to 0.997)	0.918 (0.835 to 1.000)	0.964 (0.926 to 1.001)
Unadjusted regression§	0.083 + 0.886 × TE	0.069 + 1.004 × TE	0.061 + 1.092 × TE	0.040 + 1.155 × TE
Adjusted regression‖	0.565 + 3.957 × TE	0.213 + 2.308 × TE	0.109 + 1.691 × TE	0.047 + 1.239 × TE
STE (HR)	Undefined	Undefined	0.77	0.92

* CI = confidence interval; DFS = disease-free survival; HR = hazard ratio; OS = overall survival; STE = surrogate threshold effect calculated on adjusted regression; TE = treatment effect on disease-free survival in the prediction model for treatment effect on overall survival.

† Rho represents the Spearman rank correlation coefficient between disease-free survival and overall survival.

‡ R^2 represents the coefficient of determination between treatment effect on disease-free survival and overall survival.

§ Unadjusted regression represents the linear regression prediction models for the treatment effect on overall survival from treatment effect on disease-free survival, unadjusted for the presence of estimation error in the treatment effects.

‖ Adjusted regression represents the linear regression prediction models for the treatment effect on overall survival from treatment effect on disease-free survival, adjusted for the presence of estimation error in the treatment effects.

Table 2. Observed and predicted treatment effect on overall survival based on the observed treatment effect on disease-free survival*

Trial label	Validation trials (reference)	Type of data	Observed HR_{DFS} (95% CI)	Observed HR_{OS} (95% CI)	Predicted HR_{OS} (95% PI)
A	Cirera et al. (15)	Published	0.55 (0.36 to 0.85)	0.60 (0.39 to 0.93)	0.50 (0.28 to 0.87)
B	CLASSIC (2)	Published	0.56 (0.44 to 0.72)	0.72 (0.52 to 1.00)	0.51 (0.36 to 0.73)
C	ACTS-GC (5)	IPD	0.65 (0.54 to 0.79)	0.67 (0.54 to 0.83)	0.61 (0.47 to 0.81)
D	INT-1018 (4)	IPD	0.66 (0.53 to 0.82)	0.75 (0.61 to 0.92)	0.63 (0.46 to 0.84)
E	GOIM- 9602 (14)	Published	0.88 (0.66 to 1.17)	0.91 (0.69 to 1.21)	0.89 (0.62 to 1.28)
F	GOIRC (13)	Published	0.92 (0.66 to 1.27)	0.90 (0.64 to 1.26)	0.94 (0.63 to 1.42)

* CI = confidence interval; DFS = disease-free survival; HR = hazard ratio; IPD = individual patient data; OS = overall survival; PI = prediction interval.

the hazard ratios on the basis of the events observed in each validation trial, whereas the prediction interval quantifies the uncertainty of the predicted hazard ratio for OS (without the information of OS) as a function of the observed hazard ratio for DFS. The difference between the confidence interval and the prediction interval is explained in further detail in the Supplementary Methods (available online).

Excellent agreement was noted between the observed and predicted hazard ratios for OS for two (13,14) (labeled E and F in Figure 3 and Table 2) of the four trials for which only summary data were available. The hazard ratio for OS predicted from the estimated hazard ratio for DFS after a median 5-year follow-up was lower than the observed value but still within the prediction interval for the two validation trials (2,15) (labeled A and B in Figure 3 and Table 2) for which IPD could not be obtained.

For the large Japanese trial investigating the effect of adjuvant treatment with S1 (5) (labeled C in Table 2 and Figure 4), the observed and predicted hazard ratios for OS were in reasonable agreement. For the trial (4) (labeled D in Table 2 and Figure 4) investigating the efficacy of adjuvant chemoradiation, the predicted hazard ratio for OS was also lower than the observed hazard ratio for OS, although the latter still fell within the 95% prediction interval.

Discussion

Our results show a very tight individual-level association between DFS and OS (Spearman rank correlation coefficient = 0.974; 95% CI = 0.971 to 0.976), indicating that in individual patients, DFS is highly predictive of OS. The strong correlation between DFS and OS can be partly attributed to the short time from relapse to death in gastric cancer (median of <12 months across all the included trials). Further, 16% of all the analyzed patients died without documented relapse and, therefore, had the same DFS and OS.

We also found a very high trial-level association between the effects of adjuvant chemotherapy on DFS and on OS, with R^2 being almost 1, which indicates that almost all of the variability in the treatment effects on OS can be explained by the treatment effects on DFS (Figures 1 and 3). We constructed the prediction limits around the regression line, which accounts for the fact that the hazard ratios of DFS and OS were estimated with errors. STE was found to be 0.92, thereby implying that a treatment producing an 8% or greater hazard reduction for recurrence can be expected to produce a statistically significant hazard reduction for death. STE also reflects the expected dilution of the treatment effect on OS as compared with the effect on DFS. With a reduced duration of follow-up, a stronger effect on DFS was required to predict a statistically significant benefit over OS. This is partly because of the loss of events due to the shorter period of observation. At 4 years, a hazard ratio of 0.74 was required. In case follow-up was truncated at 2 or 3 years, STE could not be estimated. A fixed follow-up period was used in this sensitivity analysis because the date of randomization was not available for all studies. In trials with an administrative censoring date common to all patients, more events would be available at the intermediate time point, resulting in more precise estimates. We also did not collect information on the treatment administered after recurrence. However, the median OS of the patients with advanced gastric cancer treated with chemotherapy was 8.7 months in the GASTRIC

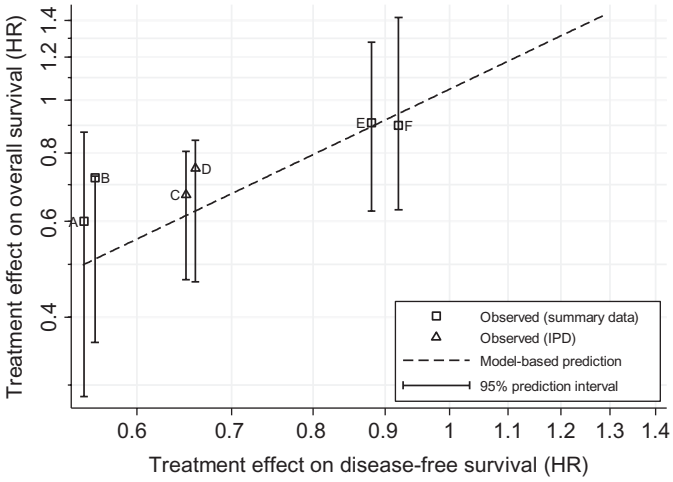

Figure 4. Observed treatment effect on disease-free survival vs predicted treatment effect on overall survival in validation trials. The **error bars** represent 95% prediction intervals. Log scale was used for the x-axis and y-axis. For the trial labels, A represents Cirera et al. (15), B represents CLASSIC (2), C represents ACTS-GC (5), D represents INT-1018 (4), E represents GOIM- 9602 (14), and F represents GOIRC (13), as shown in Table 2. HR = hazard ratio; IPD = individual patient data.

database (31), and the impact of chemotherapy on OS after relapse was not much greater than that of adjuvant chemotherapy. Thus the fact that some patients received chemotherapy on relapse is not expected to have a major impact on our findings.

In a future trial testing a new treatment for gastric cancer, interest would focus on predicting the effects on OS at some time point (eg, 5 years), having observed the effects on DFS at an earlier time point. The results presented in Table 1 suggest that the measurement of DFS at 2 years may be too early to enable an accurate prediction of OS at 5 years. With very early time points, only few DFS events are available, which may result in imprecise predictions; on the other hand, very late time points are less useful because they are closer to the evaluation of the final endpoint. Making analysis at 3 or 4 years would probably reduce the overall duration by about 15% to 30% if the accrual was short enough.

Do the present results justify the use of DFS as a surrogate for OS in resectable gastric cancer? A large proportion of relapses occurred before 3 years, and we found a strong correlation between the endpoints, both at the individual and trial levels. Similar results have led to the adoption of the 3-year DFS as a surrogate for 5-year OS in evaluating new treatments for resectable colon cancer (32). Our results are based on fewer trials and smaller sample sizes, but they include a broader range of treatment options. One may be interested in whether DFS would be a surrogate for OS for all studies independent of geography because it is well known that there exists a large heterogeneity about the prognosis between Asian and non-Asian patients. In spite of the prognostic heterogeneity between continents, there was no statistically significant heterogeneity about the treatment effects on OS and on DFS between Asia and non-Asia trials (6). In addition, the relationship between the hazard ratio for OS and the hazard ratio for DFS was also clearly consistent throughout all trials. Therefore, we believe the use of DFS as a surrogate would be independent of the geography. Moreover, we were able to use the published results of trials not included in our meta-analysis, as well as the IPD from two large trials, as two independent validation sets. The hazard ratios fell within the prediction intervals for all six available trials (Table 2). The results of the four trials with literature data only should, however, be interpreted with caution because they are based on extracted summary statistics. The relationship between the treatment effect on DFS and that on OS, as established in trials comparing adjuvant chemotherapy with surgery alone, seems to be verified for the chemoradiation trial, which implies that our results might be applicable to more general adjuvant treatments with a curative intent.

One should keep in mind the following limitations. Numerical computational issues may have slightly biased correlation estimates. Subgroup analysis based on the baseline variables, including continents, could not be performed because of the small number of trials. Similar to the case with interim analyses, follow-up after the analysis of the surrogate endpoint (DFS) is necessary to determine the OS, safety, and post-relapse outcomes as well as to document the possible impact of postrelapse treatments for advanced diseases. An important consideration is that we only investigated cytotoxic agents, and future trials investigating agents with different mechanisms of actions, such as target therapy, will require separate validation of the surrogacy relation before it is applied routinely.

In conclusion, the treatment effect on OS is largely predictable according to that on DFS; therefore, DFS can be used as a primary endpoint for further clinical trials of adjuvant chemotherapies, thus reducing the duration by 15% to 30% and the cost, depending on the planned follow-up, of these large-scale randomized trials.

References

1. Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893–2917.
2. Bang Y-J, Kim Y-W, Yang H-K, et al. Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial. *Lancet*. 2012;379(9813):315–321.
3. Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med*. 2006;355(1):11–20.
4. Macdonald JS, Smalley SR, Benedetti J, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N Engl J Med*. 2001;345(10):725–730.
5. Sasako M, Sakuramoto S, Katai H, et al. Five-year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1 versus surgery alone in stage II or III gastric cancer. *J Clin Oncol*. 2011;29(33):4387–4393.
6. The GASTRIC group. Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis. *JAMA*. 2010;303(17):1729–1737.
7. Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. *Stat Methods Med Res*. 2008;17(5):467–475.
8. Mauguen A, Pignon J-P, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. *Lancet Oncol*. 2013;14(7):619–626.
9. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000;1(1):49–67.
10. Burzykowski T, Molenberghs G, Buyse M. *The Evaluation of Surrogate Endpoints*. New York: Springer; 2006
11. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *J R Stat Soc Ser C Appl Statist*. 2001;50(4):405–422.
12. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to

progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol*. 2008;26(12):1987–1992.
13. Di Costanzo F, Gasperoni S, Manzione L, et al. Adjuvant chemotherapy in completely resected gastric cancer: a randomized phase III trial conducted by GOIRC. *J Natl Cancer Inst*. 2008;100(6):388–398.
14. De Vita F, Giuliani F, Orditura M, et al. Adjuvant chemotherapy with epirubicin, leucovorin, 5-fluorouracil and etoposide regimen in resected gastric cancer patients: a randomized phase III trial by the Gruppo Oncologico Italia Meridionale (GOIM 9602 Study). *Ann Oncol*. 2007;18(8):1354–1358.
15. Cirera L, Balil A, Batiste-Alentorn E, et al. Randomized clinical trial of adjuvant mitomycin plus tegafur in patients with resected stage III gastric cancer. *J Clin Oncol*. 1999;17(12):3810–3815.
16. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statist Med*. 1998;17(24):2815–2834.
17. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Statist*. 2006;5(3):173–186.
18. The Gastrointestinal Tumor Study Group. Controlled trial of adjuvant chemotherapy following curative resection for gastric cancer. The Gastrointestinal Tumor Study Group. *Cancer*. 1982;49(6):1116–1122.
19. Bajetta E, Buzzoni R, Mariani L, et al. Adjuvant chemotherapy in gastric cancer: 5-year results of a randomised study by the Italian Trials in Medical Oncology (ITMO) Group. *Ann Oncol*. 2002;13(2):299–307.
20. Coombes RC, Schein PS, Chilvers CE, et al. A randomized trial comparing adjuvant fluorouracil, doxorubicin, and mitomycin with no treatment in operable gastric cancer. International Collaborative Cancer Group. *J Clin Oncol*. 1990;8(8):1362–1369.
21. Bouché O, Ychou M, Burtin P, et al. Adjuvant chemotherapy with 5-fluorouracil and cisplatin compared with surgery alone for gastric cancer: 7-year results of the FFCD randomized phase III trial (8801). *Ann Oncol*. 2005;16(9):1488–1497.
22. Engstrom PF, MacIntyre JM, Douglass HO, Muggia F, Mittelman A. Combination chemotherapy of advanced colorectal cancer utilizing 5-fluorouracil, semustine, dacarbazine, vincristine, and hydroxyurea: a phase III trial by the Eastern Cooperative Oncology Group (EST: 4275). *Cancer*. 1982;49(8):1555–1560.
23. Krook JE, O'Connell MJ, Wieand HS, et al. A prospective, randomized evaluation of intensive-course 5-fluorouracil plus doxorubicin as surgical adjuvant chemotherapy for resected gastric cancer. *Cancer*. 1991;67(10):2454–2458.
24. Lise M, Nitti D, Marchet A, et al. Final results of a phase III clinical trial of adjuvant chemotherapy with the modified fluorouracil, doxorubicin, and mitomycin regimen in resectable gastric cancer. *J Clin Oncol*. 1995;13(11):2757–2763.
25. Macdonald JS, Fleming TR, Peterson RF, et al. Adjuvant chemotherapy with 5-FU, adriamycin, and mitomycin-C (FAM) versus surgery alone for patients with locally advanced gastric adenocarcinoma: a Southwest Oncology Group study. *Ann Surg Oncol*. 1995;2(6):488–494.
26. Nakajima T, Kinoshita T, Nashimoto A, et al. Randomized controlled trial of adjuvant uracil-tegafur versus surgery alone for serosa-negative, locally advanced gastric cancer. *Br J Surg*. 2007;94(12):1468–1476.
27. Nakajima T, Nashimoto A, Kitamura M, et al. Adjuvant mitomycin and fluorouracil followed by oral uracil plus tegafur in serosa-negative gastric cancer: a randomised trial. Gastric Cancer Surgical Study Group. *Lancet*. 1999;354(9175):273–277.
28. Nashimoto A, Nakajima T, Furukawa H, et al. Randomized trial of adjuvant chemotherapy with mitomycin, fluorouracil, and cytosine arabinoside followed by oral fluorouracil in serosa-negative gastric cancer: Japan Clinical Oncology Group 9206-1. *J Clin Oncol*. 2003;21(12):2282–2287.
29. Nitti D, Wils J, Dos Santos JG, et al. Randomized phase III trials of adjuvant FAMTX or FEMTX compared with surgery alone in resected gastric cancer. A combined analysis of the EORTC GI Group and the ICCG. *Ann Oncol*. 2006;17(2):262–269.
30. Tsavaris NB, Tentas K, Kosmidis P, et al. 5-Fluorouracil, epirubicin, and mitomycin C versus 5-fluorouracil, epirubicin, mitomycin C, and leucovorin in advanced gastric carcinoma. A randomized trial. *Am J Clin Oncol*. 1996;19(5):517–521.
31. The GASTRIC Group. Role of chemotherapy for advanced/recurrent gastric cancer: an individual-patient-data meta-analysis. *Eur J Cancer*. 2013;49(7):1565–1577.
32. 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 Clin Oncol*. 2005;23(34):8664–8670.

Funding

This work was partially supported by the French Institut National du Cancer (grant PHRC GASTRIC); the Clinical Research Support Unit; and the Epidemiological and Clinical Research Information Network. A meeting was supported financially by unrestricted grants from GlaxoSmithKline. The funders were not present at the meeting, were not involved in the analysis of the data, and did not comment on the present paper.

Notes

K. Oba and X. Paoletti are co–first authors; M. Buyse and T. Burzykowski are co–last authors.

The project was initiated under the auspice of the French National Cancer Institute (INCa), which served as the sponsor. The INCa centralized all the databases and provided administrative and data management support. The sponsor has not participated in the analysis and interpretation of the data, which were solely the responsibility of the writing committee. The sponsor played no part in the preparation, review, or approval of the manuscript. The conclusions may not reflect the recommendations of the INCa.

The GASTRIC Group: Secretariat: Marc Buyse, Stefan Michiels, Koji Oba, Xavier Paoletti, Philippe Rougier, Seiichiro Yamamoto, and Kenichi Nakamura. Steering Committee: Yung-Jue Bang (Seoul National University College of Medicine, Seoul, Korea); Harry Bleiberg (Brussels, Belgium); Tomasz Burzykowski (Hasselt University, Diepenbeek, Belgium); Marc Buyse (International Drug Development Institute, Louvain-la-Neuve, Belgium); Catherine Delbaldo (Hôpital Louis Mourier, Colombes, France); Stefan Michiels (Institut Gustave-Roussy, Villejuif, France); Satoshi Morita (Yokohama City University, Yokohama, Japan); Koji Oba (Hokkaido University Hospital, Sapporo, Japan); Yasuo Ohashi (University of Tokyo, Tokyo, Japan); Xavier Paoletti (Institut Curie, Paris, France); Jean-Pierre Pignon (Institut Gustave-Roussy, Villejuif, France); Philippe Rougier (University Hospital Europeen Georges Pompidou, Paris, France); Junichi Sakamoto (Tokai Central Hospital, Kakumuhara, Japan); Daniel Sargent (Mayo Clinic, Rochester, Minnesota); Mitsuru Sasako (Hyogo College of Medicine, Nishinomiya, Japan); and Eric Van Cutsem (University Hospital Gasthuisberg, Leuven, Belgium). Collaborators who supplied individual patient data: Steven Alberts, Mayo Clinic, Rochester, Minnesota; Emilio Bajetta, Department of Medical Oncology, IRCCS Fondazione Istituto Nazionale dei Tumori, Milan, Italy; Jacqueline Benedetti, SWOG Statistical Center, Seattle, Washington; Franck Bonnetain, FFCD, Dijon, France; Olivier Bouche, Department of Gastro-Intestinal Oncology, University Hospital Robert Debré, Reims, France; Paul Catalano, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; R. Charles Coombes, Medical Oncology Unit, Charing Cross Hospital, London, UK; Maria Di Bartolomeo, Department of Medical Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; Juan J. Grau, Department of Oncology, Institut Clinic de Malalties Hemato-Oncologiques of Hospital Clinic, University of Barcelona, Spain; James E. Krook, Mayo Clinic, Rochester, Minnesota; Florian Lordick, University Clinic Leipzig; University Cancer Center, Leipzig, Germany; Mario Lise, Department of Oncological and Surgical Sciences, University of Padua, Padua, Italy; John S. Mac Donald, St. Vincent's Cancer Center, New York, New York; Pierre Michel, Department of Gastro-Intestinal Oncology, Charles-Nicolle University Hospital, Rouen, France; Toshifusa Nakajima, Cancer Institute Hospital, Tokyo, Japan; Atsushi Nashimoto, Niigata Cancer Center Hospital, Niigata, Japan; Garth D. Nelson, Mayo Clinic, Rochester, Minnesota; Donato Nitti, Department of Oncological and Surgical Sciences, University of Padua, Italy; Jan Kulig, Piotr Kolodziejczak, and Tadeusz Popiela, 1st Chair of General and GI Surgery, Jagiellonian University Medical College, Krakow, Poland; Philippe Rougier, University Hospital Ambroise Paré, Boulogne,

France; Nicolas Tsavaris, Laiko University Hospital, Athens, Greece; and Mitsuru Sasako, Hyogo College of Medicine, Nishinomiya, Japan.

The GASTRIC Group thanks all the patients who took part in the trials and contributed to this research. The meta-analysis would not have been possible without the active participation of the collaborating institutions that provided trial data: Eastern Cooperative Oncology Group, European Organization of Research and Treatment of Cancer, Fédération Francophone de Cancérologie Digestive, Gastro-intestinal Tumour Study Group, International Collaborative Cancer Group, Italian Trials in Medical Oncology, Japan Clinical Oncology Group, North Central Cancer Treatment Group, South West Oncology Group, Hospital Clinic Villarroel of Barcelona, Metaxa Cancer Hospital, Jagiellonian of Pireus University, and Medical College of Cracow. We thank Nicolas Thamavong for his help in data management.

The first draft of this manuscript was developed at a meeting of investigators in Sapporo, Japan, September 24–28, 2010. The meeting was supported financially by unrestricted grants from GlaxoSmithKline.

Affiliations of authors: Translational Research and Clinical Trial Center, Hokkaido University Hospital, Sapporo, Japan (KO); Biostatistics dpt, INSEM U900, Institut Curie, Paris, France (XP); Mayo Clinic, Rochester, MN (SA, DS); Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea (Y-JB); SWOG Statistical Center, Seattle, WA (JB); Brussels, Belgium (HB); Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard School of Public Health, Boston, MA (PC); University Clinic Leipzig, University Cancer Center, Leipzig, Germany (FL); Biostatistics and Epidemiology department, Institut Gustave-Roussy, Universite Paris XI, Villejuif, France (SMi, J-PP); Department of Biostatistics, Yokohama City University, Yokohama, Japan (SMo); Department of Biostatistics, School of Public Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan (YO); Gastroenterology department, University Hospital Europeen Georges Pompidou, Paris, France (PR); Department of Surgery, Hyogo College of Medicine, Nishinomiya, Japan (MS); Tokai Central Hospital, Kakumuhara, Japan (JS); Department of Gastroenterology and Gastrointestinal Oncology, National Cancer Center Hospital East, Kashiwa, Japan (KS); Digestive Oncology Unit, University Hospital Gasthuisberg, Leuven, Belgium (EVC); International Drug Development Institute, Louvain-la-Neuve, Belgium (MB, TB); I-BioStat, Hasselt University, Diepenbeek, Belgium (MB, TB).