The Oncologist®

Lung Cancer

Validation of Progression-Free Survival as a Surrogate Endpoint for Overall Survival in Malignant Mesothelioma: Analysis of Cancer and Leukemia Group B and North Central Cancer Treatment Group (Alliance) Trials

Xiaofei Wang,[a,b] Xiaoyi Wang,[a] Lydia Hodgson,[b] Stephen L. George,[a] Daniel J. Sargent,[c] Nate R. Foster,[c] Apar Kishor Ganti,[d] Thomas E. Stinchcombe,[e] Jeffrey Crawford,[e] Robert Kratzke,[f] Alex A. Adjei,[g] Hedy L. Kindler,[h] Everett E. Vokes,[h] Herbert Pang[a,i]

[a]Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina, USA; [b]Alliance Statistics and Data Center, Duke University, Durham, North Carolina, USA; [c]Alliance Statistics and Data Center, Mayo Clinic, Rochester, Minnesota, USA; [d]Department of Internal Medicine, Veterans Affairs Nebraska Western Iowa Health Care System and University of Nebraska Medical Center, Omaha, Nebraska, USA; [e]Department of Medicine, Duke University, Durham, North Carolina, USA; [f]Department of Medicine, University of Minnesota, Minneapolis, Minnesota, USA; [g]Department of Medical Oncology, Mayo Clinic, Rochester, Minnesota, USA; [h]Department of Medicine, University of Chicago, Chicago, Illinois, USA; [i]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, People's Republic of China

Disclosures of potential conflicts of interest may be found at the end of this article.

Key Words. Malignant mesothelioma • Surrogate endpoint • Progression-free survival • Overall survival • Risk factors

Abstract

Purpose. The aim of this study was to investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in malignant mesothelioma.

Materials and Methods. Individual data were collected from 15 Cancer and Leukemia Group B (615 patients) and 2 North Central Cancer Treatment Group (101 patients) phase II trials. The effects of 5 risk factors for OS and PFS, including age, histology, performance status (PS), white blood cell count, and European Organisation for Research and Treatment of Cancer (EORTC) risk score, were used in the analysis. Individual-level surrogacy was assessed by Kendall's tau through a Clayton bivariate Copula survival (CBCS) model. Summary-level surrogacy was evaluated via the association between logarithms of the hazard ratio (log HR)—log HR_{OS} and log HR_{PFS}—measured in R^2 from a weighted least-square (WLS) regression model and the CBCS model.

Results. The median PFS for all patients was 3.0 months (95% confidence interval [CI], 2.8–3.5 months) and the median OS was 7.2 months (95% CI, 6.5–8.0 months). Moderate correlations between PFS and OS were observed across all risk factors at the individual level, with Kendall's tau ranging from 0.46 to 0.47. The summary-level surrogacy varied among risk factors. The Copula R^2 ranged from 0.51 for PS to 0.78 for histology. The WLS R^2 ranged from 0.26 for EORTC and PS to 0.67 for age.

Conclusions. The analyses demonstrated low to moderate individual-level surrogacy between PFS and OS. At the summary level, the surrogacy between PFS and OS varied significantly across different risk factors. With a short postprogression survival and a moderate correlation between PFS and OS, there is no evidence that PFS is a valid surrogate endpoint for OS in malignant mesothelioma. ***The Oncologist*** 2017;22:189–198

Implications for Practice: For better disease management and for more efficient clinical trial designs, it is important to know if progression-free survival (PFS) is a good surrogate endpoint for overall survival in malignant mesothelioma. With a relatively large database of 17 phase II trials and 716 patients from Cancer and Leukemia Group B and North Central Cancer Treatment Group, we conducted statistical analyses and found that there is no evidence to suggest that PFS is a valid surrogate endpoint for OS for malignant mesothelioma. Future research work is needed to find alternative surrogate endpoints for OS.

Introduction

Mesothelioma is a rare cancer that originates in the mesothelium and is often caused by exposure to asbestos [1]. The annual number of new cases in the United States is 2,000–3,000 [2]. Mesothelioma is frequently diagnosed at an advanced stage; therefore, patients have a poor prognosis. The median survival after surgery is approximately 1 year [1]. Most

Correspondence: Xiaofei Wang, Ph.D., Duke University, 2424 Erwin Road, Suite 1102, Durham, North Carolina 27705, USA. Telephone: 919-681-5406; e-mail: xiaofei.wang@duke.edu Received March 23, 2016; accepted for publication August 18, 2016; published Online First on February 10, 2017. 1083-7159/2016/$20.00/0 http://dx.doi.org/10.1634/theoncologist.2016-0121

Table 1. Summary of 17 mesothelioma trials from Cancer and Leukemia Group B and North Central Cancer Treatment Group

Study	Reference	Treatment	Relapsed disease	Patients (*n*)	Median PFS (95% CI), mo	Median OS (95% CI), mo
C8435 arm A	Chahinian et al. (1993) [11]	Cisplatin + mitomycin	None	39	3.65 (2.66–7.13)	7.95 (5.39–11.24)
C8435 arm B	Chahinian et al. (1993) [11]	Cisplatin + doxorubicin	None	39	5.85 (2.76–7.95)	9.27 (5.03–11.17)
C8638	Vogelzang et al. (1990) [12]	Carboplatin	None	41	2.76 (1.91–3.84)	7.10 (5.78–8.97)
C8833	Vogelzang et al. (1997) [13]	Dihydro-5	None	44	2.27 (1.51–3.38)	7.31 (4.37–10.12)
C8933	Vogelzang et al. (1994) [14]	Trimetrexate	None	52	3.19 (2.04–5.06)	8.38 (6.14–11.60)
C9031	Samuels et al. (1998) [15]	Dihydro-5-azacytidine + cisplatin	None	36	3.15 (1.87–4.50)	6.46 (4.34–10.15)
C9131 arm A	Kindler et al. (1999) [16]	Edatrexate	None	20	5.60 (2.10,13.83)	10.88 (5.32–18.0)
C9131 arm B	Kindler et al. (1999) [16]	Edatrexate + leucovorin	None	40	3.45 (2.04–3.71)	6.65 (5.55–7.85)
C9234	Vogelzang et al. (1999) [17]	Paclitaxel + G-CSF	None	35	3.19 (1.58–4.14)	5.13 (3.88–7.03)
C9530	Kindler et al. (2001) [18]	Gemcitabine	None	17	1.68 (0.99–3.58)	4.67 (1.48–10.71)
C9631	Kosty et al. (2001) [19]	Doxorubicin + dexrazoxane + GM-CSF	None	10	2.76 (0.33–3.19)	4.65 (0.33–15.61)
C9733	Kindler et al. (2005) [20]	Irinotecan	None	28	2.73 (1.74–2.92)	9.30 (4.11–13.24)
C39807	Otterson et al. (2004) [21]	Capecitabine	None	27	2.43 (1.41–3.98)	4.86 (2.56–9.69)
C30101	Govindan et al. (2005) [22]	Gefitinib	None	43	1.71 (1.41–3.58)	6.80 (3.09–8.41)
C30107	Jahan et al. (2012) [23]	Vatalanib	None	47	4.11 (2.00–5.32)	9.99 (6.24–14.39)
C30307	Dubey et al. (2010) [24]	Sorafenib	Some patients	51	3.58 (2.07–5.2)	9.00 (4.21–14.32)
C30601	Dudek et al. (2012) [25]	Dasatinib	All patients	46	2.10 (1.71–3.61)	6.01 (3.98–7.66)
N0021	Okuno et al. (2008) [26]	Gemcitabine + epirubicin	None	67	4.73 (2.69–6.28)	7.98 (5.62–9.36)
N0623	Molina et al. (2009) [27]	Pazopanib	None	34	4.17 (1.71–5.95)	11.53 (5.26–18.20)

Abbreviations: CI, confidence interval; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; OS, overall survival; PFS, progression-free survival.

patients die in a short period after disease progression. The combination of a platinum agent (carboplatin or cisplatin) and an antifolate agent (pemetrexed or raltitrexed) is standard therapy for patients with advanced disease [3–5].

Overall survival (OS), the time from study registration or randomization to death, is the most important clinically meaningful endpoint for evaluating drug efficacy in oncology clinical trials. However, OS may not always be the optimal primary endpoint for efficacy assessment. It often requires a longer follow-up period, and any potential effect of first-line therapy may be attenuated by effective sequential therapies, making it difficult to evaluate the true treatment effect of the first-line therapy. When newer agents display promising early efficacy, many trials allow the patients on standard therapy crossover to the investigational agent after progression. In such cases, the use of a surrogate endpoint, such as progression-free survival (PFS), in place of OS is an appealing option. In this study, we assessed the capability of PFS as a surrogate endpoint for OS for malignant mesothelioma based on data from the clinical trials conducted by Cancer and Leukemia Group B (CALGB) and North Central Cancer Treatment Group (NCCTG) trials. CALGB and NCCTG are now part of the Alliance for Clinical Trials in Oncology.

A surrogate endpoint, such as PFS, can be used to evaluate treatment effect much like an established endpoint, such as OS, in evaluating an experimental treatment. In clinical cancer research, there is a general interest in validating or invalidating PFS as a surrogate endpoint for OS [6, 7]. Rather than pursuing a strict definition of surrogacy, such as the Prentice criteria [8], we adopted a more practical definition on surrogacy [9] and the corresponding statistical methods to validate the PFS surrogacy [10]. Specifically, surrogacy evaluation can take place at three different levels: individual level (correlation between PFS and OS within a patient), trial level (correlation between the treatment effect on PFS and that on OS), and summary level (correlation between the effect of risk factor on PFS and that on OS). When the data are from randomized clinical trials, the surrogacy of PFS as a potential endpoint to replace OS in evaluating new treatment is often based on the first two measures: individual-level and trial-level surrogacy. Indeed, such approach has been successfully used to validate PFS as a surrogate endpoint for OS in colorectal cancer [6], in which treatment effects based on PFS and OS were used for evaluation.

The primary goal of the current study was to evaluate the utility of PFS as a surrogate endpoint for OS in mesothelioma by using clinical trials data from CALGB and NCCTG. For rare disease, such as malignant mesothelioma, clinical trials data are mostly from small single-arm trials; we are not able to estimate the trial-level surrogacy, which directly answers whether a treatment-induced change in PFS predicts a treatment-induced change in OS. In such situations, we have to rely on the individual-level correlation, including the summary-level correlation based on risk groups, for surrogacy assessment. We feel

that different approaches or measures for individual-level correlation are warranted in surrogacy evaluation, although our approach is not for the purpose of estimating the trial-level correlation.

Materials and Methods

Trials and Data

CALGB and NCCTG conducted 17 trials (15 CALGB and 2 NCCTG) on malignant mesothelioma between 1984 and 2009 (Table 1) [11–27]. Follow-up data reported before February 2014 were included. Of the 17 trials, 15 trials were single-arm phase II trials and 2 were two-arm phase II randomized trials. Each arm of the two-arm trials was considered as an independent trial in our analysis. Because mesothelioma is a relatively rare disease, the number of patients enrolled in each trial was small, ranging from 10 to 67. Each participant signed an institutional review board-approved, protocol-specific informed consent in accordance with federal and institutional guidelines. Data were collected by CALGB and NCCTG statistics and data centers.

The following data were extracted for each trial: gender, age at registration, Eastern Cooperative Oncology Group (ECOG) performance status (PS), previous radiation status, previous chemotherapy status, previous surgery status, weight loss, mesothelioma histology, white blood cell count (WBC), date of last observation (or date of death, if patient died), survival status, date of progression, progression status, baseline primary tumor, baseline nodal involvement, and baseline distant metastasis.

PFS Surrogacy Evaluation with Single-Arm Trial Data

In this analysis, PFS is defined as the time from study registration or randomization to disease progression or death from any cause, whichever comes first. The common approach of validating PFS surrogacy is to use data from large randomized clinical trials in which patients were randomly assigned to the same (or similar) two treatment arms. Most CALGB and NCCTG mesothelioma trials are single-arm trials, and the treatments varied from trial to trial. Hence, the common approach of evaluating PFS surrogacy on treatment effect is no longer applicable. Risk factors, such as ECOG PS and histology, are known prognostic factors for PFS and OS in mesothelioma [28, 29]. Like treatment effect, these risk factors affect both PFS and OS. We postulated that these risk factors affect PFS and OS in a manner similar to that of treatment effect (e.g., pemetrexed plus cisplatin vs. cisplatin alone). Therefore, for the data of mostly single-arm trials, it is reasonable to use the risk factors, such as PS and histology, to substitute treatment effect for PFS surrogacy evaluation. To verify this idea, we examined the effect of treatments (experimental vs. control) on PFS and OS and the effect of PS (1/2 vs. 0) by using the extensive-stage small cell lung cancer (ESCLC) data assembled for another study [30], in which both treatment effect and risk factors were available. As seen in the supplemental online data (Appendix A), PFS surrogacy at the individual level and the summary level is very similar for the PS effect and the treatment effect in the ESCLC dataset. Although the summary-level analysis based on the risk factor effect is consistent with the trial-level analysis based on the treatment effect in the example of small cell lung cancer, there is a fundamental difference between the surrogacy valuation based on true treatment effect and that based on risk factor effect.

To investigate PFS surrogacy induced by risk factors on PFS and OS, we focused on the following risk factors identified in two well-known prognostic models for mesothelioma [28, 29]: ECOG PS [31], histology, age, and WBC. All patients were classified into binary risk groups: PS ($\geq$1 or 0), age ($\geq$70 or $<$70 years), histology (nonepithelial or epithelial), and WBC ($>$0.83 or $\leq 0.83 \times 10^9$/L). For the surrogacy analysis, we also considered European Organisation for Research and Treatment of Cancer (EORTC) risk score [28] as a risk factor; this is a composite score of several risk factors, including WBC, PS, and histology. On the basis of EORTC risk score, patients were classified into a good prognosis group ($\leq$1.27) and a poor prognosis group ($>$1.27).

Statistical Analysis

Survival Analysis

The distributions of PFS and OS were estimated by using the Kaplan-Meier method. Estimation procedures and hypothesis tests were stratified by the classification criteria for trials (Fig. 1). The effects of risk factors on PFS and OS were estimated as logarithms of the hazard ratio (log HR), respectively: log HR_{OS} and log HR_{PFS} log, through Cox proportional hazards model with the risk factor of interest as a single predictor. The log HR is approximately equal to the risk reduction for small effects. The HR estimates were used for assessing the agreement between PFS effect and OS effect at the summary level. There was no intention to conclude whether any specific risk factor is statistically significant at either the summary level or overall. For this reason, the issue of multiplicity was not addressed.

Surrogacy Criteria Analysis

To evaluate whether PFS is a valid surrogate endpoint for OS, we aimed to demonstrate that (a) PFS was strongly associated with OS (individual-level surrogacy) and (b) the treatment effect (in this case the risk factor effect) for PFS and for OS was strongly associated (summary-level surrogacy). The association of log HR_{OS} and log HR_{PFS} log was assessed to evaluate the summary-level surrogacy. A weighted least-square regression for log HR_{OS} and log HR_{PFS} log was performed to evaluate the relationship of OS and PFS, with weights equal to the sample size of the trial [6, 7]. Another approach used a bivariate survival model derived by Buyse et al. [10], in which both summary-level and individual-level surrogacies between PFS and OS were estimated. The bivariate survival model was constructed from a Clayton Copula with Weibull marginal distribution. For individual-level surrogacy, the association parameter δ was transformed into a scale parameter, Kendall's tau, which belongs to the interval $[-1, 1]$. A value of tau near 1 suggests a strong positive association between PFS and OS, whereas a value near -1 indicates a negative association. Finally, a leave-one-out cross-validation was performed to assess the surrogacy between PFS and OS [32]. Specifically, we calculated the mean squared prediction error (MSPE) by using the leave-one-out cross-validation procedure by averaging the square of the difference between observed HR_{OS} and predicted HR_{OS} over all trials. The percentage of patients with one risk factor missing ranged from 0% for age to 14% for EORTC risk score; the patients with missing risk factors were excluded from the

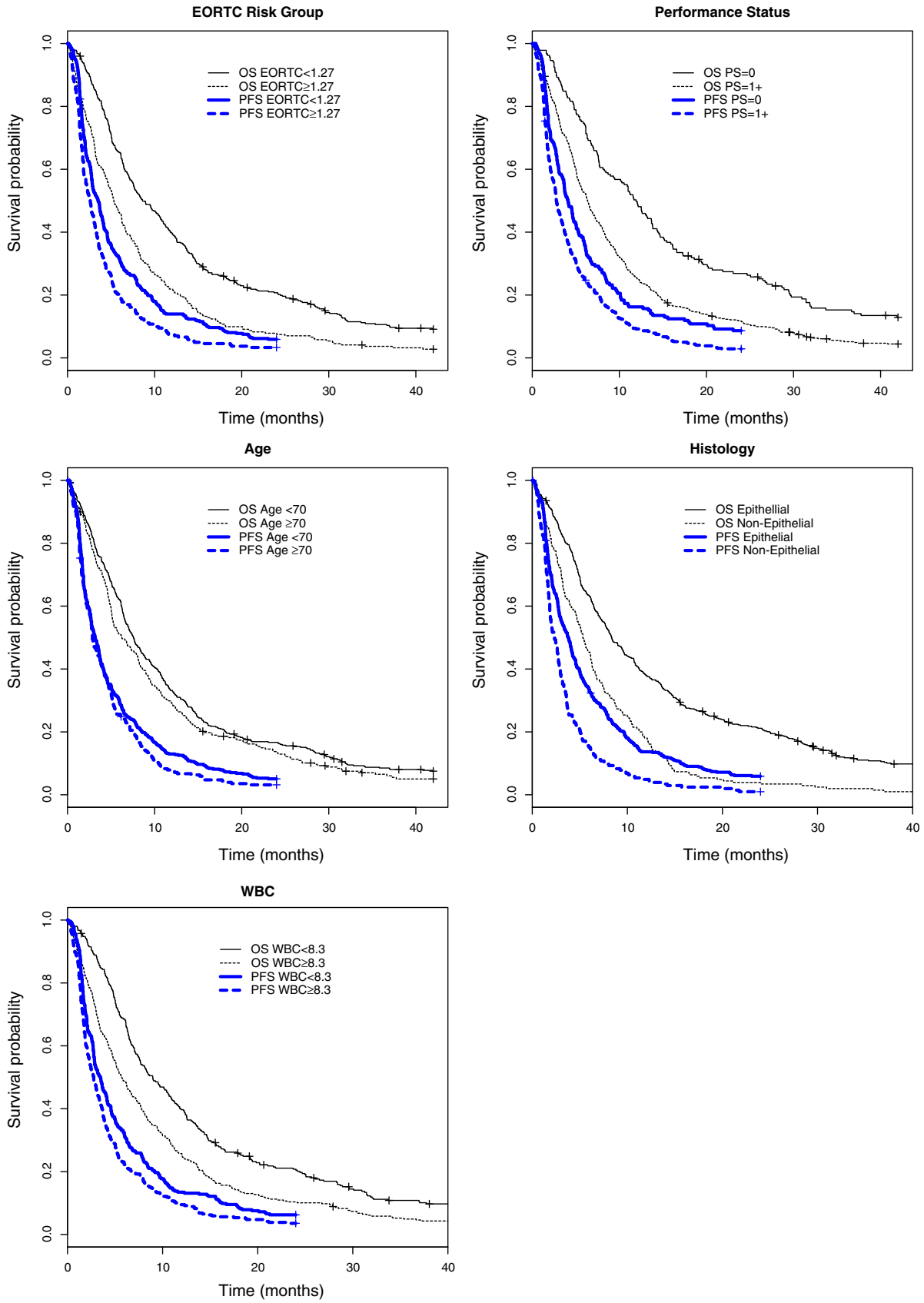

Figure 1. Kaplan-Meier curves of OS and PFS.

Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; OS, overall survival; PFS, progression-free survival; WBC, white blood cell count.

analysis of that particular risk factor. Approximately 15% patients had received prior chemotherapy at the enrollment of these trials. The primary analysis included all patients, but a sensitivity analysis including only patients who received no prior chemotherapy was also conducted. All statistical analyses were done in SAS 9.4 (SAS, Cary, NC, USA, http://www.sas.com) and R 3.3.2 (R core team, Vienna, Austria, http://www.R-project.org).

Results

Patient Characteristics and Outcomes

The data were from the 716 patients enrolled in the 17 mesothelioma clinical trials. As shown in Table 2, the median age was 65 years (range, 23–92 years); 81% of patients were male. For the entire cohort, the median follow-up time was 42 months. The median PFS and OS were 3.0 months (95% confidence interval [CI], 2.8–3.5 months) and 7.2 months (95% CI, 6.5–8.0 months), respectively. Two thirds (67.2%) of the patients had epithelioid histology and 25.8% of the patients had a PS score of 0. EORTC risk score was calculated as follows: risk score = 0.55 (if WBC $>8.3 \times 10^9$/L), 0.60 (if performance status ≥ 1), 0.52 (if histologic diagnosis is probable or possible), 0.67 (if histologic subtype is sarcomatous), and 0.60 (if gender is male). A total of 616 (86%) patients had all risk factors observed, and we were able to compute the EORTC risk score; 39.3% of these patients were classified into the poor prognosis group (>1.27) and 60.7% were in the good prognosis group (≤ 1.27).

Effects of Risk Factors on PFS and OS

Figure 1 shows the PFS and OS curves of all patients for five risk factors, including EORTC risk group (poor or good prognosis), PS (≥ 1 or 0), age group (≥ 70 or <70 years), histology (nonepithelial or epithelial) and WBC (≥ 8.3 or $<8.3 \times 10^9$/L). Each plot represents Kaplan-Meier curves of PFS and OS by different levels of a risk factor. For all patients, the HRs for EORTC risk group (poor and good prognosis) were 1.36 for PFS (95% CI, 1.15–1.6) and 1.67 for OS (95% CI, 1.41–1.97). The median PFS and median OS were 2.6 months and 5.3 months for the poor prognosis group, respectively, and 3.5 months and 8.6 months for the good prognosis group. A clear separation of curves is evident in each plot for both endpoints, and the benefit ordering is consistent between OS and PFS, which suggests an association between OS and PFS in malignant mesothelioma.

We also calculated the summary-level PFS and OS HRs for each risk group. Supplemental online Figure 1 displays the estimated HRs and their 95% CIs by each trial for EORTC risk group. Similar forest plots (supplemental online Fig. B2–B5) for other risk groups can also be found in supplemental online Appendix B. C9631 was excluded from these plots because with only 10 patients its HR for risk factors was inestimable or unreasonably big. These plots indicate moderate agreement between PFS HRs and OS HRs in terms of direction and effect size across risk factors. A separate set of analyses was conducted to include only the 607 patients who did not receive prior chemotherapy at enrollment. Supplemental online Figure 1 shows similar Kaplan-Meier plots for PFS and OS for patients receiving no prior chemotherapy.

Table 2. Summary of baseline patient characteristics

Variable	Patients ($n = 716$), n (%)
Gender	
Male	582 (81.3)
Female	134 (18.7)
Age group	
<70 yr	456 (63.7)
≥ 70 yr	260 (36.3)
EORTC risk group[#]	
Good	374 (52.2)
Poor	242 (33.8)
Missing	100 (14.0)
Histology	
Epithelial	481 (67.2)
Nonepithelial	205 (28.6)
Missing	30 (4.2)
PS	
0	185 (25.8)
≥ 1	528 (73.7)
Missing	3 (0.4)
White blood cell count	
$\geq 8.3 \times 10^9$/L	306 (42.7)
$<8.3 \times 10^9$/L	338 (47.2)
Missing	72 (10.1)
Weight loss	
$<5\%$	267 (37.3)
$\geq 5\%$	177 (24.7)
Missing	272 (38.0)
Previous surgery	
Yes	463 (64.7)
No	214 (29.9)
Missing	39 (5.5)
Previous radiation	
Yes	66 (9.2)
No	646 (90.2)
Missing	4 (0.6)
Previous chemotherapy	
Yes	107 (14.9)
No	605 (84.5)
Missing	4 (0.6)

[a]Good prognosis if EORTC ≤ 1.27 and poor prognosis if EORTC >1.27.
Abbreviations: CI, confidence interval; EORTC, European Organisation for Research and Treatment of Cancer; OS, overall survival; PFS, progression-free survival; PS, performance status.

Individual-Level Surrogacy

The Clayton bivariate Copula survival (CBCS) model was used to evaluate individual-level surrogacy. Across all trials, Kendall's tau ranged from 0.45 to 0.47 for all risk factors, indicating that PFS and OS were only moderately correlated at the individual-level (Table 3). Supplemental online Table C1 shows similar results for the patients receiving no prior chemotherapy.

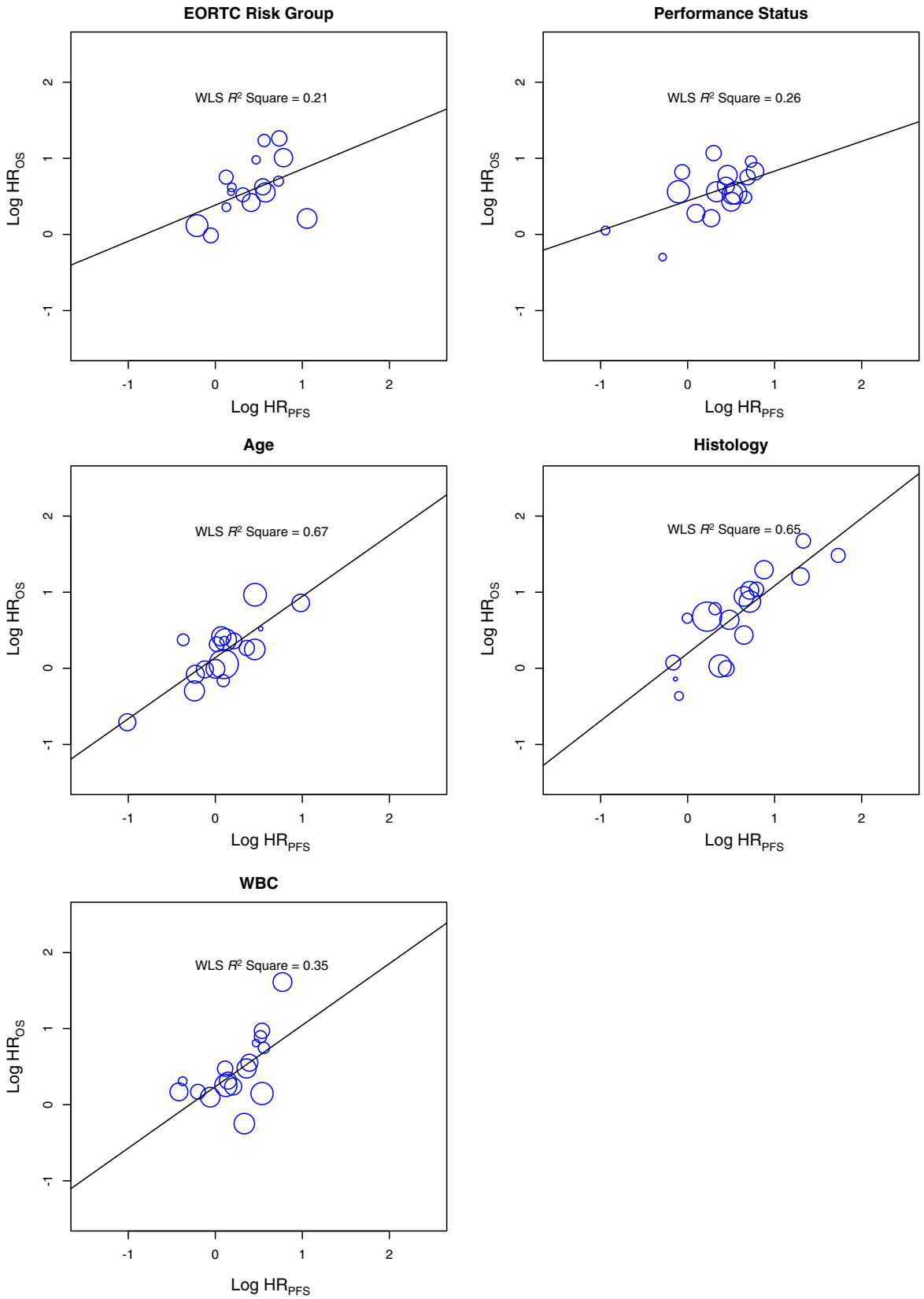

Figure 2. WLS regression of log HR_{OS} on log HR_{PFS}.

Abbreviations: log HR_{OS}, logarithm of the hazard ratio for overall survival; log HR_{PFS}, logarithm of the hazard ratio for progression-free survival; WBC, white blood cell count; WLS, weighted least-square.

Table 3. Results from Clayton bivariate Copula survival model

Variable	Patients (*n*)	Copula R^2 (95% CI)[a]	Tau[b] (95% CI)
EORTC[c] (poor, good prognosis)	242, 374	0.69 (0.45–0.93)	0.45 (0.40–0.50)
PS (≥1, 0)	185, 528	0.61 (0.34–0.89)	0.45 (0.40–0.49)
Age (≥70 yr, <70 yr)	260, 456	0.51 (0.19–0.82)	0.45 (0.40–0.49)
Histology (nonepithelial, epithelial)	205, 481	0.78 (0.61–0.96)	0.46 (0.41–0.50)
WBC (>8.3, ≤8.3 × 10^9/L)	338, 306	0.76 (0.56–0.95)	0.47 (0.42–0.51)

[a]Copula R^2 is a measure for summary-level surrogacy based on Clayton bivariate Copula survival model.
[b]Kendall's tau is a measure for individual-level surrogacy.
[c]Good prognosis if EORTC ≤1.27 and poor prognosis if EORTC >1.27.
Abbreviations: CI, confidence interval; EORTC, European Organisation for Research and Treatment of Cancer; PS, performance status; WBC, white blood cell count.

Table 4. Results from weighted least-square model

Variable	β_0 (95% CI)[a]	β_1 (95% CI)[b]	WLS R^2[c]	MSE[d]	MSPE[e]
EORTC (poor, good prognosis)	0.35 (0.12–0.58)	0.51 (0.08–0.95)	0.26	0.09	0.14
PS (≥1, 0)	0.44 (0.28–0.61)	0.39 (0.06–0.72)	0.26	0.07	0.09
Age (≥70 yr, <70 yr)	0.14 (0.03–0.26)	0.80 (0.53–1.08)	0.67	0.05	0.07
Histology (nonepithelial, epithelial)	0.20 (−0.05 to 0.44)	0.89 (0.57–1.21)	0.65	0.10	0.12
WBC (>8.3, ≤8.3 × 10^9/L)	0.24 (0.01–0.46)	0.81 (0.25–1.37)	0.35	0.11	0.15

C9631 was removed from WLS analysis because with only 10 patients it made the hazard ratios for several risk factors inestimable or unreasonably big.
[a]Intercept estimated from WLS regression model.
[b]Slope estimated form WLS regression model.
[c]WLS R^2 is a measure for summary-level surrogacy based on WLS model.
[d]MSE of β_1 based on WLS model.
[e]MSPE of β_1 based on leave-one-out validation.
Abbreviations: CI, confidence interval; EORTC, European Organisation for Research and Treatment of Cancer; MSE, mean square error; MSPE, mean square prediction error; PS, performance status; WBC, white blood cell count. WLS, weighted least-square.

Summary-Level Surrogacy

Two measures were used to assess summary-level surrogacy. According to the modeling of the CBCS model, the summary-level surrogacy was moderate across all the risk groups, with Copula R^2 values ranging from 0.51 for age to 0.78 for histology (Table 3).

For each risk group, a weighted least-squares (WLS) regression model can be fitted with summary-level HRs and weighted with the size of the trial (Fig. 2). For example, the linear regression equation for EORTC risk group was as follows (Fig. 2, Table 4):

$$\log HR(OS) = 0.35 + 0.51 * \log HR(PFS)$$

This model indicates that the risk increase associated with poor prognosis group was approximately 49% lower on OS than on PFS. The closer the slope of the regression line is to 1, the better the agreement between the effect of that risk group on PFS and OS.

The WLS R^2 ranged from 0.26 to 0.67; the lowest was for EORTC and PS and the highest was for age (Table 4). In the WLS analysis, we excluded one obvious outlier that significantly affected the goodness of fit if the data point were included. C9631 was removed from WLS analysis because with only 10 patients it made the HRs for several risk factors inestimable or unreasonably big. The shrinkage of PFS surrogacy for OS was further evaluated by using leave-one-out cross-validation. The values of MSPE for all risk factors are generally bigger than the corresponding mean square error values. Such shrinkage of prediction accuracy on new data is expected [32]. The lowest MSPE was for age (0.07) and highest was for WBC (0.15) (Table 4). Similar results were observed for summary-level surrogacy for the patients receiving no prior chemotherapy, as seen in supplemental online Figure C2, Table C1, and Table C2.

Summary-Level Surrogacy Based on Artificially Paired Treatment Arms

As seen in Table 1, two studies (C8435 and C9131) are randomized phase II trials and the cohort consists of 138 patients. The small sample size prevented us from carrying out the analysis on summary-level surrogacy. For a purely exploratory exercise, we computed summary-level correlation by creating artificial pairs of treatment arms. Specifically, the two arms of the randomized trial each formed a pair. C9631 was excluded because it had only 10 patients. The rest of 14 single-arm trials formed 7 pairs based on the closeness between the means of the EORTC risk scores. For each pair, the trial/arm with better PFS was coded as the experimental arm, and the trial/arm with worse PFS was coded as the control arm. On the basis of on the bivariate Copula model, the individual-level tau is 0.46 and the trial-level correlation Copula R^2 is 0.77. On the basis of the weighted least square model, the summary-level correlation WLS R^2 is 0.52. These estimates again suggest a moderate PFS surrogacy for OS when data from single arm trials were paired into treatment groups by the proximity of their EROTC scores.

Discussion

The goal of this project was to evaluate whether PFS can be used as a surrogate endpoint for OS in malignant mesothelioma clinical trials. Individual patient data from a total of 716

patients enrolled in 17 CALGB and NCCTG phase II trials were used to evaluate the surrogacy of PFS for OS at both the patient and summary levels. Our analysis suggests that PFS is associated with OS, but the measures on both individual-level and summary-level surrogacy are not strong enough to allow the use of PFS as a surrogate endpoint for OS, regardless of which risk factor was considered.

In the individual-level analysis, the concordance between PFS and OS, as measured by Kendall's tau values, ranged from 0.45 to 0.47, suggesting a low to moderate correlation. In the summary-level analysis, the proportion of OS effect explainable by the corresponding PFS effect was measured by Copula R^2 and WLS R^2. Although in the bivariate Copula model, the Copula R^2 was moderately high for some risk factors, such as histology (0.78) and WBC group (0.76), it was low for the remainder, ranging from 0.51 to 0.69. In the WLS analysis, there was also a weak correlation between PFS and OS across all the trials; R^2 values ranged from very low (0.26) for EORTC risk group and PS to moderate (0.67) for age. As an exploratory exercise, the trial-level surrogacy was estimated by the artificial randomized trials data created by pairing single-arm trials on the means of the EORTC risk scores; again, the PFS surrogacy was moderate and similar to those estimated from the risk factor effect approach.

Our analyses have several limitations. Most of the trials we analyzed contained only one treatment arm, and the treatment varied from trial to trial. The PFS surrogacy for OS has usually been evaluated with multiple trials, with the same two treatment regimens across these trials [6, 7]. One can argue that the effect of baseline prognostic factors, such as PS, on survival endpoints is different from the treatment effect on the same endpoints and that these baseline factors are not the causes of survival difference. Although the use of the effect of risk factors as a substitute for treatment effect in evaluating PFS surrogacy is supported by the analysis of small cell lung cancer data reported in supplemental online data (Appendix A) and this approach has been used by a different team of investigators for similar analysis [33, 34], the validity of this approach needs further confirmation from both methodological and applied research.

In addition, most of these trials are small, with sample sizes ranging from 10 to 67, causing bigger errors associated with summary-level estimates. Moreover, none of the 17 trials contained pemetrexed. Active agents, such as pemetrexed, may have an effect on the observed correlation between PFS HRs and OS HRs, and the direction and the magnitude of the impact are not assessable with the existing data. These limitations may have mischaracterized the real surrogacy relation between PFS and OS. However, because mesothelioma is a rare disease, it is unlikely that there will be many large clinical trials for this population. This analysis includes a large number of patients enrolled on multiple cooperative group trials, and hence the results likely outweigh the limitations.

The trial-level surrogacy cannot be estimated from single-arm trial data because no patients of the same trial were assigned to different treatment arms. To evaluate PFS surrogacy with single-arm trial data, we had to resort to the summary-level surrogacy, which is obtained by replacing treatment arms with levels of risk factors and then applying the standard surrogacy analysis. The concordance of PFS surrogacy evaluated with the risk factor effect and the treatment effect in the small cell lung cancer analysis suggests that the surrogacy tends to take place at multiple levels. However, we need to emphasize that the summary-level correlation is fundamentally different from the trial-level correlation. This statement holds for the exploratory exercise in which artificial single-arm treatment trial was paired into pseudo-randomized treatment groups. In other words, the creation of artificial "treatment groups" is another way of estimating the summary-level correlation, and it does not inform the true trial-level correlation. The summary-level correlation carries no information on trial-level surrogacy in the conventional sense.

Meanwhile, when the risk factor effect is used for surrogacy analysis on single-arm trials data, the individual correlation tau estimated from the bivariate Copula model is related but, in general, distinct from the individual-level surrogacy obtained from randomized trials data. For similar reasons, the square error and the MSPE in Table 4 may have a different interpretation when these statistics are estimated by using randomized trials data. The analysis includes all 17 mesothelioma trials conducted in CALGB and NCCTG, and most the investigative therapies were negative. This may dampen the generalizability of our findings.

Additional issues for validating PFS as a surrogate endpoint in clinical trials have been discussed by other authors [35–38], such as PFS definition varying among studies and assessment bias in open-label studies. Many of the trials performed in mesothelioma have been small phase II trials with investigator review of response and PFS. These issues are not unique to our analysis but make it more difficult to validate PFS as a surrogate endpoint for OS in malignant mesothelioma.

Conclusion

More mesothelioma patients are expected to receive maintenance or second-line treatments, and there is a need to demonstrate that PFS or other short-term endpoints can be used as a valid surrogate endpoint for OS. Because of these limitations of our database and the analysis approach, we shall refrain from a definitive conclusion that PFS is not a valid surrogate endpoint for OS in future mesothelioma trials. A definitive answer to the surrogacy question in this disease requires further investigation on the data from multiple large randomized trials, with disease progression confirmed by an independent central review [39, 40]. Hence, pending more definitive studies, we cannot recommend using PFS as a surrogate endpoint for OS in malignant mesothelioma trials.

Acknowledgments

We thank the following study chairs for their contribution in conducting the malignant mesothelioma trials: Dr. A.P. Chahinian (C8435), Dr. N.J. Vogelzang (C8638, C8833, C8933, C9234), Dr. B.L. Samuels (C9031), Dr. H.L. Kindler (C9131, C9530, C9733), Dr. M.P. Kosty (C9631), Dr. G.A. Otterson (C39807), Dr R. Govindan (C30101), Dr. T. Jahan (C30107), Dr. S. Dubey (C30307), Dr A.Z. Dudek (C30601), Dr S.H. Okuno (N0021), and Dr. J.R. Molina (N0623).

This study was supported by NCI cooperative group research grants U10CA033601, U10CA180882, U10CA180821, U10CA180838, U10CA180857, U10CA180866, U10CA180836,

P01CA142538; NIA R21 grant 5R21AG042894; VA Career Development Award; and the PCFB grant.

Author Contributions

Conception/Design: Xiaofei Wang, Stephen L. George, Daniel J. Sargent, Apar Kishor Ganti, Thomas E. Stinchcombe, Everett E. Vokes, Herbert Pang

Collection and/or assembly of data: Xiaofei Wang, Lydia Hodgson, Robert Kratzke, Hedy L. Kindler, Everett E. Vokes, Herbert Pang

Data analysis and interpretation: Xiaofei Wang, Xiaoyi Wang, Stephen L. George, Daniel J. Sargent, Nate R. Foster, Apar Kishor Ganti, Thomas E. Stinchcombe, Jeffrey Crawford, Everett E. Vokes, Herbert Pang

Manuscript writing: Xiaofei Wang, Xiaoyi Wang, Daniel J. Sargent, Apar Kishor Ganti, Thomas E. Stinchcombe, Jeffrey Crawford, Alex A. Adjei, Everett E. Vokes, Herbert Pang

Final approval of manuscript: Xiaofei Wang, Xiaoyi Wang, Lydia Hodgson, Stephen L. George, Daniel J. Sargent, Nate R. Foster, Apar Kishor Ganti, Thomas E. Stinchcombe, Jeffrey Crawford, Robert Kratzke, Alex A. Adjei, Hedy L. Kindler, Everett E. Vokes, Herbert Pang

Disclosures

Apar K. Ganti: Pfizer, NewLink Genetics, Janssen Pharmaceuticals, Merck, Amgen (RF); **Jeffrey Crawford**: Merck, Novartis, Pfizer, Celgene, G1 Therapeutics, Janssen, Merrimack, Roche (CA); Amgen, AstraZeneca, Bayer, Clovis, GTx, Morphotek (RF); **Everett E. Vokes**: Amgen, Boehringer-Ingelheim, Celgene, Eli Lilly, Genentech, Merck, Synta, VentiRx (C/A). The other authors indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

References

1. Jaklitsch MT, Grondin SC, Sugarbaker DJ. Treatment of malignant mesothelioma. World J Surg 2001;25:210–217.

2. American Lung Association. Understanding Mesothelioma. 2015. http://www.lung.org/lung-health-and-diseases/lung-disease-lookup/mesothelioma/

3. Santoro A, O'Brien ME, Stahel RA et al. Pemetrexed plus cisplatin or pemetrexed plus carboplatin for chemonaive patients with malignant pleural mesothelioma: Results of the International Expanded Access Program. J Thorac Oncol 2008;3:756–763.

4. van Meerbeeck JP, Gaafar R, Manegold C et al. Randomized phase III study of cisplatin with or without raltitrexed in patients with malignant pleural mesothelioma: An intergroup study of the European Organisation for Research and Treatment of Cancer Lung Cancer Group and the National Cancer Institute of Canada. J Clin Oncol 2005;23:6881–6889.

5. Vogelzang NJ, Rusthoven JJ, Symanowski J et al. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. J Clin Oncol 2003;21:2636–2644.

6. Buyse M, Burzykowski T, Carroll K et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 2007;25:5218–5224.

7. 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 2007;26:1987–1992.

8. Prentice RL. Surrogate endpoints in clinical trials: Definition and operational criteria. Stat Med 1989;8:431–440.

9. Buyse M, Molenberghs G. Criteria for the validation of surrogate endpoints in randomized experiments. Biometrics 1998;54:1014–1229.

10. Buyse M, Molenberghs G, Burzykowski T et al. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics 2000;1:49–67.

11. Chahinian AP, Antman K, Goutsou M et al. Randomized phase II trial of cisplatin with mitomycin or doxorubicin for malignant mesothelioma by the Cancer and Leukemia Group B. J Clin Oncol 1993;11:1559–1565.

12. Vogelzang NJ, Goutsou M, Corson JM et al. Carboplatin in malignant mesothelioma: A phase II study of the Cancer and Leukemia Group B. Cancer Chemother Pharmacol 1990;27:239–242.

13. Vogelzang NJ, Weissman LB, Herndon JE 2nd et al. Trimetrexate in malignant mesothelioma: A Cancer and Leukemia Group B Phase II study. J Clin Oncol 1994;12:1436–1442.

14. Vogelzang NJ, Herndon JE 2nd, Cirrincione C et al. Dihydro-5-azacytidine in malignant mesothelioma. A phase II trial demonstrating activity accompanied by cardiac toxicity. Cancer and Leukemia Group B. Cancer 1997;79:2237–2242.

15. Samuels BL, Herndon JE 2nd, Harmon DC et al. Dihydro-5-azacytidine and cisplatin in the treatment of malignant mesothelioma: A phase II study by the Cancer and Leukemia Group B. Cancer 1998;82:1578–1584.

16. Kindler HL, Belani CP, Herndon JE 2nd et al. Edatrexate (10-ethyl-deaza-aminopterin) (NSC #626715) with or without leucovorin rescue for malignant mesothelioma. Sequential phase II trials by the cancer and leukemia group B. Cancer 1999;86:1985–1991.

17. Vogelzang NJ, Herndon JE 2nd, Miller A et al. High-dose paclitaxel plus G-CSF for malignant mesothelioma: CALGB phase II study 9234. Ann Oncol 1999;10:597–600.

18. Kindler HL, Millard F, Herndon JE 2nd et al. Gemcitabine for malignant mesothelioma: A phase II trial by the Cancer and Leukemia Group B. Lung Cancer 2001;31:311–317.

19. Kosty MP, Herndon JE 2nd, Vogelzang NJ et al. High-dose doxorubicin, dexrazoxane, and GM-CSF in malignant mesothelioma: A phase II study-Cancer and Leukemia Group B 9631. Lung Cancer 2001;34:289–295.

20. Kindler HL, Herndon JE, Zhang C et al. Irinotecan for malignant mesothelioma A phase II trial by the Cancer and Leukemia Group B. Lung Cancer 2005;48:423–428.

21. Otterson GA, Herndon JE 2nd, Watson D et al. Capecitabine in malignant mesothelioma: A phase II trial by the Cancer and Leukemia Group B (39807). Lung Cancer 2004;44:251–259.

22. Govindan R, Kratzke RA, Herndon JE 2nd et al. Gefitinib in patients with malignant mesothelioma: A phase II study by the Cancer and Leukemia Group B. Clin Cancer Res 2005;11:2300–2304.

23. Jahan T, Gu L, Kratzke R et al. Vatalanib in malignant mesothelioma: A phase II trial by the Cancer and Leukemia Group B (CALGB 30107). Lung Cancer 2012;76:393–396.

24. Dubey S, Janne PA, Krug L et al. A phase II study of sorafenib in malignant mesothelioma: Results of Cancer and Leukemia Group B 30307. J Thorac Oncol 2010;5:1655–1661.

25. Dudek AZ, Pang H, Kratzke RA et al. Phase II study of dasatinib in patients with previously treated malignant mesothelioma (Cancer and Leukemia Group B 30601): A brief report. J Thorac Oncol 2012;7:755–759.

26. Okuno SH, Delaune R, Sloan JA et al. A phase 2 study of gemcitabine and epirubicin for the treatment of pleural mesothelioma: A North Central Cancer Treatment Study, N0021. Cancer 2008;112:1772–1779.

27. Molina JR, Mandrekar SJ, Ziegler A, et al. A phase II study of the VEGFR inhibitor pazopanib (P) in patients with newly diagnosed or relapsed malignant pleural mesothelioma (MPM): An NCCTG study. J Thorac Oncol 2009;(suppl 1):S456.

28. Curran D, Sahmoud T, Therasse P et al. Prognostic factors in patients with pleural mesothelioma: The European Organization for Research and Treatment of Cancer experience. J Clin Oncol 1998;16:145–152.

29. Herndon JE, Green MR, Chahinian AP et al. Factors predictive of survival among 337 patients with mesothelioma treated between 1984 and 1994 by the Cancer and Leukemia Group B. Chest 1998;113:723–731.

30. Foster NR, Renfro LA, Schild SE et al. Multitrial evaluation of progression-free survival as a surrogate end point for overall survival in first-line extensive-stage small-cell lung cancer. J Thorac Oncol 2015;10:1099–1106.

31. Oken MM, Creech RH, Tormey DC et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982;5:649–655.

32. Geisser S. Predictive Inference. New York: Chapman and Hall, 1993.

33. Lee-Ying RM, Speers C, Gresham G et al Population-based assessment of surrogate endpoints (SE) in stage I-III colon (CCa) and rectal cancers (RCa). J Clin Oncol 2015;33: abstr 3538.

34. Lee-Ying RM, Yin Y, Lim HJ et al. Population-based assessment of disease-free survival (DFS) as a surrogate endpoint (SE) of 5-year overall survival (5Y OS) in stage I-III rectal (RC) and colon cancer (CC). J Clin Oncol 2016;34(suppl): abstr 3602.

35. U.S. Food and Drug Administration. Guidance for industry: Clinical trial endpoints for the approval of cancer drugs and biologics. 2007. Available at: http://www.fda.gov/downloads/Drugs/. . ./Guidances/ucm071590.pdf. Accessed December 3, 2016.

36. Booth CM, Eisenhauer EA: Progression-free survival: Meaningful or simply measurable? J Clin Oncol 2012;30:1030–1033.

37. 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;23:8664–8670.

38. Burzykowski T, Molenberghs G, Buyse M et al. Validation of surrogate end points in multiple randomized clinical trials with failure time end points. J Royal Stat Soc Series C (Appl Stat) 2001;50: 405–422.

39. Panageas KS, Ben-Porat L, Dickler MN et al. When you look matters: The effect of assessment schedule on progression-free survival. J Natl Cancer Inst 2007;99:428–432.

40. Dancey JE, Dodd LE, Ford R et al. Recommendations for the assessment of progression in randomised cancer treatment trials. Eur J Cancer 2009;45:281–289.

See http://www.TheOncologist.com for supplemental material available online.

For Further Reading:

Jing Ai, James P. Stevenson. Current Issues in Malignant Pleural Mesothelioma Evaluation and Management. ***The Oncologist*** 2014;19:975–984.

Implications for Practice:

Although uncommon, malignant pleural mesothelioma (MPM) is being diagnosed at an increasing rate worldwide due to continued workplace exposure in developing countries to asbestos and other potentially carcinogenic inhaled silicates. This article emphasizes the need for multidisciplinary evaluation at diagnosis to identify appropriate candidates for multimodality therapy and to optimize survival outcomes for this deadly disease. A growing body of data suggests that lung-sparing extended pleurectomy is the option of choice for most patients who are surgical candidates. Insights into altered molecular pathways and the immunology of MPM have led to clinical trials of novel drugs.