Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer

Peer-reviewed author version

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Evaluation of Tumor Response, Disease Control, Progression-Free Survival, and Time to Progression As Potential Surrogate End Points in Metastatic Breast Cancer

Tomasz Burzykowski, Marc Buyse, Martine J. Piccart-Gebhart, George Sledge, James Carmichael, Hans-Joachim Lück, John R. Mackey, Jean-Marc Nabholtz, Robert Paridaens, Laura Biganzoli, Jacek Jassem, Marijke Bontenbal, Jacques Bonneterre, Stephen Chan, Gul Atalay Basaran, and Patrick Therasse

From the Hasselt University, Diepenbeek; International Drug Development Institute, Louvain-la-Neuve; University Hospital Gasthuisberg, Leuven; Institut Jules Bordet; and European Organization for Research and Treatment of Cancer, Brussels, Belgium; Indiana University-Purdue University, Indianapolis, IN; Astra Zeneca, Macclesfield; and Nottingham City Hospital, United Kingdom; Medizinische Hochschule, Hannover, Germany; University of Alberta, Edmonton, Canada; Breast Cancer Research Institute–La Prandie, Valojoulx; and Centre Oscar Lambret, Lille, France; Hospital of Prato, Italy; Medical University of Gdansk, Poland; Erasmus MC, University Medical Center, Rotterdam, the Netherlands; and Marmara University Hospital, Istanbul, Turkey.

Corresponding author: Tomasz Burzykowski, Hasselt University, Center for Statistics, Agoralaan, Bldg D, 3590 Diepenbeek, Belgium; e-mail: tomasz.burzykowski@uhasselt.be.

ABSTRACT

Background
Overall survival (OS) can be observed only after prolonged follow-up, and any potential effect of first-line therapies on OS may be confounded by the effects of subsequent therapy. We investigated whether tumor response, disease control, progression-free survival (PFS), or time to progression (TTP) could be considered a valid surrogate for OS to assess the benefits of first-line therapies for patients with metastatic breast cancer.

Patients and Methods
Individual patient data were collected on 3,953 patients in 11 randomized trials that compared an anthracycline (alone or in combination) with a taxane (alone or in combination with an anthracycline). Surrogacy was assessed through the correlation between the end points as well as through the correlation between the treatment effects on the end points.

Results
Tumor response (survival odds ratio [OR], 6.2; CI, 5.3 to 7.0) and disease control (survival OR, 5.5; CI, 4.8 to 6.3) were strongly associated with OS. PFS (rank correlation coefficient [ρ], 0.688; CI, 0.686 to 0.690) and TTP (ρ, 0.682; CI, 0.680 to 0.684) were moderately associated with OS. Response log ORs were strongly correlated with PFS log hazard ratios (ρ, 0.96; CI, 0.73 to 1.19). Response and disease control log ORs and PFS and TTP log hazard ratios were poorly correlated with log hazard ratios for OS, but the confidence limits of ρ were too wide to be informative.

Conclusion
No end point could be demonstrated as a good surrogate for OS in these trials. Tumor response may be an acceptable surrogate for PFS.

J Clin Oncol 26:000-000. © 2008 by American Society of Clinical Oncology

INTRODUCTION

Overall survival (OS) is viewed as the end point of choice to assess the efficacy of new treatments in advanced breast cancer. However, the analysis of OS requires prolonged follow-up of all patients, thereby delaying the evaluation of highly promising new therapies that will often find their most cost-effective use in the adjuvant setting. In addition, any potential effect of first-line therapies on OS may be diluted (but possibly not eliminated) by the effects of subsequent therapy with agents that have activity against breast cancer. It is, therefore, important to investigate whether other clinical end points, such as tumor response, disease control, progression-free survival (PFS), or time

to progression (TTP), could replace OS as the primary end point in randomized trials for the first-line chemotherapy of patients with advanced breast cancer.

To address this question, we utilized randomized trials of first-line treatment for advanced breast cancer; three trials compared a single-agent taxane with a single-agent anthracycline, and eight further trials explored the value of combining taxanes and anthracyclines.[1-11] Individual patient data from these 11 trials were collected and reanalyzed. The results of this meta-analysis and a more detailed description of the trials are published in a companion paper (Piccart-Gebhart et al, manuscript submitted for publication). In this article, we quantify the relationships of tumor response, disease control,

balt5/zlj-jco/zlj-jco/zlj01208/zlj6842-08a | kroenink | S=3 | 3/11/08 | 10:36 | 4/Color Figure(s): F1-4 | Art: 08407 | Input-NMV

Burzykowski et al

PFS, and TTP as candidates for surrogate end points for OS by using data from the same set of trials.

PATIENTS AND METHODS

Trials

The 11 randomized trials either compared anthracycline-taxane combination regimens versus anthracycline-based regimens or compared single-agent anthracycline versus single-agent taxane regimens for the first-line treatment of metastatic breast cancer (Piccart-Gebhart et al, manuscript submitted for publication; Table 1).

T1
AQ: A
AQ: B

Data

The following data items were collected for all individual patients included in all trials: center, random assignment date, date of last observation (or date of death, if patient died), survival status, cause of death, tumor response (according to the WHO criteria[12]), date of tumor response, date of progression, progression status, number of organs involved at entry, visceral disease at entry (defined as predominantly visceral, if available, or as lung and/or liver), and estrogen receptor status at entry (defined as positive or negative according to each center's or trial's policy). PFS was calculated from the date of random assignment to the date of disease progression or death as a result of any cause; observations of patients who were alive without progression at the end of their follow-up (including those who received subsequent anticancer therapy in the absence of documented disease progression) were considered censored observations. TTP was calculated from the date of random assignment to the date of disease progression or death caused by the disease; observations of patients who were alive without progression at the end of their follow-up (including those who received subsequent anticancer therapy in the absence of documented disease progression) and patients who died as a result of causes not related to the disease were considered censored observations. Survival was calculated from date of random assignment to the date of death as a result of any cause; observations of patients who were alive at the end of their follow-up were considered censored observations. Because the information on the date of tumor response or disease stabilization was not available for all trials, disease control was defined as the achievement of a complete response, a partial response, or stable disease that had a TTP greater than 6 months.

Validation Strategy

The validity of surrogate end points (tumor response, disease control, PFS, and TTP as surrogates for survival and tumor response as a surrogate for PFS) was investigated through measures of association between the end points and through a joint model to estimate trial-specific treatment effects on these end points.[13] The log odds ratio (OR) was used to quantify the effect of treatment on response or on disease control, and the log hazard ratio (HR) was used to quantify the effect of treatment on PFS, TTP, and OS. The association of treatment effects was analyzed by using linear regression. The correlation coefficient was adjusted for the presence of the estimation error in treatment effects by using an error-in-variables linear regression model.[14]

The association between end points (ie, individual patient outcomes) is related to the prediction of the end point of choice (eg, OS) from the surrogate end point (eg, PFS) for an individual patient. The stronger the association, the better the precision of the prediction is and the more valid the surrogate is at the individual patient level, as it allows for a reliable replacement of the end point of choice by the surrogate. However, this association merely captures the prognostic aspect of a surrogate.

The association between treatment effects (differences in average outcomes for treated and control patients) is related to the prediction of the treatment effect on the end point of choice (eg, log HR for OS) from the effect on surrogate (eg, log HR for PFS) at the trial level. Again, the stronger the association, the more valid the surrogate is at the trial level, as it allows for a more precise prediction of the treatment effect. In fact, it is this prediction that is of most interest: one would like to decide on a new treatment by looking at the effect of the treatment on the surrogate instead of the effect on the true end

Table 1. Clinical Trials Included in the Meta-Analysis		
Study and Arm	Treatment	No. of Patients
Combination trials		3,034
Paclitaxel		1,763
UKCCCR AB01[1]		705
Control arm	E 75/C 600	
Treatment arm	E 75/P 200	
AGO[2]		516
Control arm	E 60/C 600	
Treatment arm	E 60/P 175	
EORTC 10961[3]		275
Control arm	A 60/C 600	
Treatment arm	A 60/P 175	
CCEI Paclitaxel BCSG[4]		267
Control arm	F 500/A 50/C 500	
Treatment arm	A 50/P 220	
Docetaxel		1,271
Tax 307 study group[5]		484
Control arm	F 500/A 50/C 500	
Treatment arm	D 75/A 50/C 500	
Tax 306 study group[6]		429
Control arm	A 60/C 600	
Treatment arm	A 50/D 75	
CCC Netherlands[7]		216
Control arm	F 500/A 50/C 500	
Treatment arm	A 50/D 75	
French trial[8]		142
Control arm	F 500/E 75/C 500	
Treatment arm	E 75/D 75	
Single agent trials		919
Paclitaxel		821
ECOG E1193[9]		490
Control arm	A 60	
Treatment arm	P l,175	
EORTC 10923[10]		331
Control arm	A 75	
Treatment arm	P 200	
Docetaxel		98
Tax 303 study group[11]*		98
Control arm	A 75	
Treatment arm	D 100	
All trials		3,953

NOTE. All doses are in mg/m².
Abbreviations: UKCCCR, United Kingdom Committee for Cancer Clinical Research; E, epirubicin; C, cyclophosphamide; P, paclitaxel (Taxol); AGO, Arbeitsgemeinschaft Gynaekologische Onkologie; EORTC, European Organisation for Research and Treatment of Cancer; A, adriamycin; CCEI Paclitaxel BCSG, Central Europe & Israel Paclitaxel Breast Cancer Study Group; F, fluorouracil; D, docetaxel (Taxotere); CCC, Comprehensive Cancer Centre; ECOG, Eastern Cooperative Oncology Group.
*The remaining patients of this trial received the drugs as second-line therapy and, therefore, were not included in the analysis.

point, because the former might be observed earlier or in more patients than the latter. It should be stressed that the trial level association between treatment effects does not simply follow from the individual level association between end points.[13]

All analyses were intention-to-treat. All confidence intervals were calculated for a two-sided probability coverage of 95%.

Tumor Response and Disease Control as Surrogate End Points

A joint model that was based on the Plackett copula was used to estimate trial-specific treatment effects on tumor response and survival by using a

balt5/zlj-jco/zlj-jco/zlj01208/zlj6842-08a | kroenink | S=3 | 3/11/08 | 10:36 | 4/Color Figure(s): F1-4 | Art: 08407 | Input-NMV

Surrogate End Points in Metastatic Breast Cancer

marginal proportional odds model for tumor response and a marginal Weibull model for survival.[15] The association between response and survival was quantified through the survival OR, an assumed constant over time, and was defined as the odds of surviving for responders divided by the odds of surviving nonresponders. The association between disease control and survival was quantified in a similar way.

To eliminate length-biased sampling from the analyses of the tumor response, we used the landmark method.[16] This method consists of ignoring responses that occur later than an arbitrary landmark time and deaths that occur before that time. We varied landmark times between 1 to 6 months to assess the sensitivity of the results to the choice of the landmark. A similar approach was adopted for the analyses of disease control.

PFS and TTP as Surrogate End Points

A joint model that was based on the Hougaard copula was used to estimate trial-specific treatment effects on PFS or TTP and survival by using marginal proportional-hazards models with normally distributed, random trial–specific treatment effects for the TTP and survival.[17] The association between PFS or TTP and survival was quantified through the Spearman rank correlation coefficient (ρ).

AQ: C

RESULTS

Treatment Effects

Table 2 shows the treatment effects on the clinical outcomes of interest: response, disease control rates, ORs, median time, and HR for survival, PFS, and TTP. Note that the effects for PFS and TTP are similar. This is because only 393 patients (9.9%) did not experience disease progression.

T2

Association Between End Points

The survival OR for responders versus nonresponders was 2.9 (CI, 2.6 to 3.3), which indicated that a patient who responded was about three times more likely to survive beyond any arbitrary time than one who did not respond. The value of the OR remained greater than 1.9 (and the lower CI limit greater than 1.6) for landmark times between 1 to 6 months.

Table 2. Treatment Effects for OS, PFS, TTP, Tumor Response, and Disease Stabilization

Study	OS Median (months)	OS HR (T/A)	OS 95% CI	PFS Median (months)	PFS HR (T/A)	PFS 95% CI	TTP Median (months)	TTP HR (T/A)	TTP 95% CI	Response Rate (%)	Response OR (T/A)	Response 95% CI	Disease Stabilization Rate (%)	Disease Stabilization OR (T/A)	Disease Stabilization 95% CI
UKCCR		1.02	0.87 to 1.19		1.06	0.92 to 1.24		1.06	0.92 to 1.24		1.49	1.11 to 2.02		1.33	0.96 to 1.87
Arm A	13.9			7.1			7.08			53			70		
Arm T	12.9			7.0			7.03			63			76		
AGO		1.16	0.94 to 1.43		0.94	0.78 to 1.13		0.94	0.78 to 1.13		1.64	1.16 to 2.33		1.77	1.22 to 2.55
Arm A	22.1			7.4			7.4			39			59		
Arm T	19.1			9.2			9.2			51			72		
EORTC 10961		1.12	0.82 to 1.53		0.95	0.74 to 1.20		0.95	0.74 to 1.20		1.17	0.73 to 1.89		1.01	0.61 to 1.66
Arm A	23.7			6.0			5.8			54			65		
Arm T	21.4			6.0			6.0			58			65		
CCEI		0.69	0.51 to 0.93		0.73	0.56 to 0.95		0.72	0.55 to 0.95		1.57	0.96 to 2.56		1.67	0.98 to 2.84
Arm A	18.3			6.6			6.2			54			65		
Arm T	22.3			8.3			8.3			65			75		
Tax307		0.96	0.78 to 1.19		0.96	0.79 to 1.17		0.96	0.79 to 1.16		1.44	1.01 to 2.07		1.17	0.81 to 1.68
Arm A	22.0			7.2			7.2			41			59		
Arm T	20.6			7.6			7.6			50			63		
Tax306		0.89	0.72 to 1.11		0.88	0.72 to 1.08		0.89	0.72 to 1.09		1.68	1.14 to 2.46		1.23	0.82 to 1.83
Arm A	21.7			7.9			8.1			46			65		
Arm T	22.5			8.6			8.7			59			69		
CCC		0.70	0.52 to 0.94		0.67	0.51 to 0.88		0.67	0.51 to 0.88		2.29	1.33 to 3.96		1.90	1.06 to 3.44
Arm A	16.2			6.7			6.7			37			63		
Arm T	22.7			8.0			8.0			58			76		
French		0.79	0.50 to 1.25		0.90	0.62 to 1.31		0.85	0.57 to 1.28		3.01	1.52 to 5.98		3.01	1.51 to 5.98
Arm A	28.0			5.5			5.5			32			32		
Arm T	34.0			5.9			5.9			58			59		
ECOG		0.94	0.78 to 1.14		0.97	0.81 to 1.16		0.97	0.81 to 1.17		0.97	0.67 to 1.41		0.98	0.69 to 1.40
Arm A	18.9			6.4			6.7			34			53		
Arm T	21.7			6.3			6.3			34			53		
EORTC 10923		1.08	0.86 to 1.36		1.69	1.35 to 2.11		1.69	1.35 to 2.11		0.49	0.31 to 0.79		0.40	0.25 to 0.62
Arm A	18.1			7.4			7.4			41			62		
Arm T	16.0			4.1			4.2			25			40		
Tax303		1.18	0.68 to 2.03		1.02	0.65 to 1.60		1.02	0.65 to 1.60		1.01	0.45 to 2.24		0.72	0.31 to 1.67
Arm A	21.7			7.6			7.6			49			67		
Arm T	16.2			5.7			5.7			49			60		

Abbreviations: OS, overall survival; PFS, progression-free survival; TTP, time to progression; HR, hazard ratio; T, taxane; A, anthracycline; OR, odds ratio; UKCCR, United Kingdom Committee for Cancer Clinical Research; AGO, Arbeitsgemeinschaft Gynaekologische Onkologie; EORTC, European Organisation for Research and Treatment of Cancer; CCC, Comprehensive Cancer Centre; ECOG, Eastern Cooperative Oncology Group.

The survival OR for patients with disease control versus other patients was 5.5 (CI, 4.8 to 6.3), which indicated that a patient with disease control was about five times more likely to survive beyond any arbitrary time than a patient without the control. The value of the OR remained greater than 2.7 (and the lower CI limit greater than 2.3) for landmark times between 1 to 6 months.

The ρ between PFS and survival was 0.688 (CI, 0.686 to 0.690), which indicated a moderate correlation between these end points. The ρ between TTP and survival was 0.682 (CI, 0.680 to 0.684).

In an additional analysis, the PFS OR for responders versus nonresponders was 6.2 (CI, 5.3 to 7.0), which indicated that a responding patient was about six times more likely to survive free of progression beyond any arbitrary time than a nonresponding patient. The value of the OR remained greater than 4 (and the lower CI limit greater than 3.7) for landmark times between 1 to 6 months.

Association Between Treatment Effects

The ρ between the treatment effects on response and on survival was 0.57 (CI, −0.31 to 1.44), which indicated a loose and imprecisely estimated association between these two end points (Fig 1). In the landmark analysis, the ρ varied slightly between 0.54 and 0.63, depending on the landmark time (data not shown).

The ρ between the treatment effects on disease control and on survival was 0.47 (CI, −0.46 to 1.40), which indicated again a loose and imprecisely estimated association between treatment effects on these two end points (Fig 2). In the landmark analysis, the ρ varied between 0.39 and 0.55, depending on the landmark time (data not shown).

The ρ between the treatment effects on PFS and survival was 0.48 (CI, −0.34 to 1.30; Fig 3). For the treatment effects on TTP and survival, the ρ was 0.49 CI, −0.32 to 1.30; plot not shown). As with the result obtained for tumor response and disease control, both estimates indicated a weak and imprecisely estimated association between treatment effects for PFS and TTP and, conversely, for survival.

In an additional analysis, the ρ between the treatment effects on response and on PFS was 0.96 (CI, 0.73 to 1.19), which
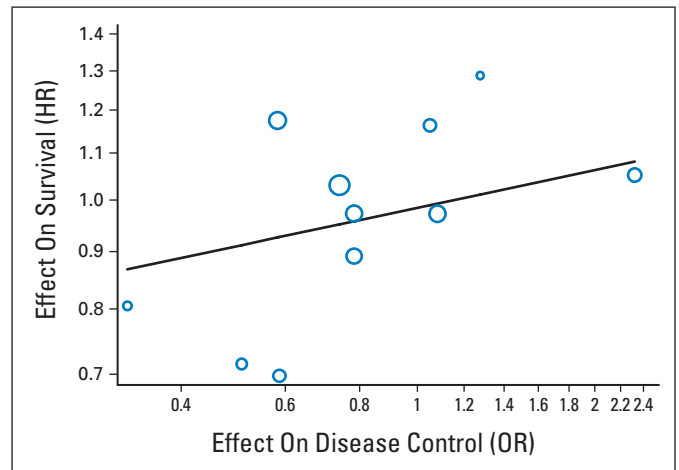

Fig 2. Correlation between treatment effects on disease control and on survival. Each circle represents a trial, and the surface area of the circle is proportional to the size of the corresponding trial.

indicated a remarkably tight association between these effects (Appendix Fig A1, online only). The equation of the regression line is as follows: log HR = 0.10 + 0.50 × log OR, which predicted a 50% increase in the PFS log HR for every unit increase in the log OR. The results remained virtually unchanged in the landmark analysis, irrespective of the choice of the landmark (data not shown).

DISCUSSION

Because of the number of new promising treatments arriving in phase III trials for advanced breast cancer, there is an urgent need to identify end points that are more sensitive and more rapidly observed than OS.[18] Before such earlier end points are adopted in clinical practice, they must be shown to be acceptable surrogates for the more meaningful clinical end points they are supposed to replace.
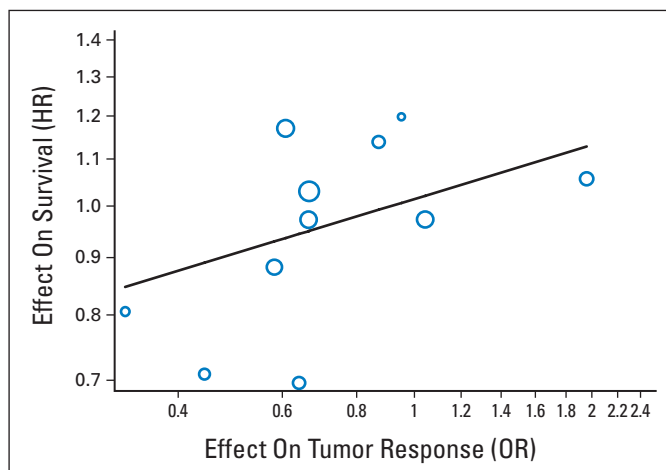

Fig 1. Correlation between treatment effects on response and on survival. Each circle represents a trial, and the surface area of the circle is proportional to the size of the corresponding trial.
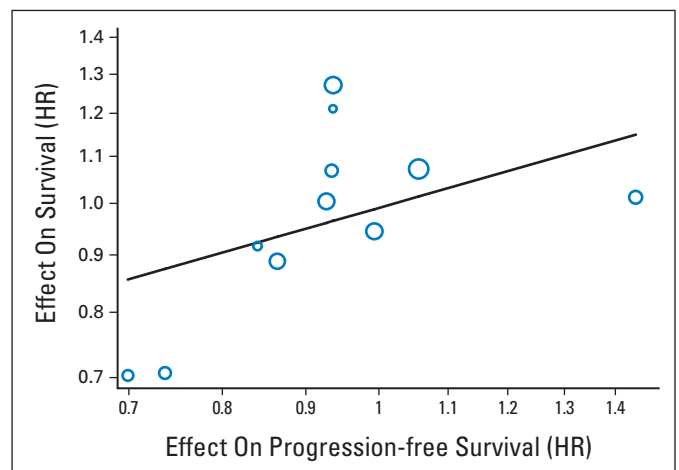

Fig 3. Correlation between treatment effects on progression-free survival and on survival. Each circle represents a trial, and the surface area of the circle is proportional to the size of the corresponding trial.

balt5/zlj-jco/zlj-jco/zlj01208/zlj6842-08a | kroenink | S=3 | 3/11/08 | 10:36 | 4/Color Figure(s): F1-4 | Art: 08407 | Input-NMV

Surrogate End Points in Metastatic Breast Cancer

Some authors have suggested, on the basis of results published in the literature, that PFS or even tumor response could be considered surrogates for survival in advanced breast cancer. Hackshaw et al[19] looked at 13 first-line randomized trials that had one FAC (fluorouracil, adriamycin, cyclophosphamide) or FEC (fluorouracil, epirubicin, cyclophosphamide) treatment arm. They used the naïve statistical approach of simply correlating treatment effects on two of these end points, by using summary statistics and by making no allowance either for the correlation between the end points or for the errors in the estimated treatment effects. This approach overestimated the ρ between the treatment effects on the end points considered and is likely to overstate the quality of an assumed surrogacy between these end points. They found a ρ of 0.66 between the treatment effects on response and survival and of 0.87 between the treatment effects on PFS and survival. In contrast, the corresponding ρ data in our analysis were 0.57 and 0.48, respectively, with so much uncertainty in both cases that a zero correlation could not be excluded. The uncertainty in our analyses may be result from the minimal effect of taxanes on OS.[1-11] This observation might lend support to the claim that, to validate a surrogate for survival, the treatment must have some effect both on the surrogate and on survival itself in at least some trials.[20-21] In addition, actual validation analyses may require a range of treatment effects for the correlation approach we advocate here to be informative.

AQ: D

Bruzzi et al[22] used a different set of randomized trials that compared a standard FEC regimen with a dose-intensified FEC regimen.[22] They collected individual patient data on 2,126 patients who were randomly assigned in 10 trials. By using a statistical approach close in spirit to ours, they showed that tumor response is a highly significant predictor of survival but that the treatment effects on response poorly predict the treatment effects on survival. Although our results agree with theirs on both counts, our conclusions differ. Bruzzi et al[22] claim that response is a potential surrogate end point for survival, but we feel that the poor prediction of the treatment effects on survival from the treatment effects on response calls for a more balanced interpretation. Although we cannot exclude that tumor response could be a surrogate for survival in a different set of trials, our data offer no evidence for this claim. Again, the lack of a survival benefit in our meta-analysis may have limited our ability to validate response as a surrogate for survival. However, a meta-analysis of randomized trials in advance colorectal cancer, in which treatment had a significant effect on both response and survival, also failed to validate response as a surrogate for survival.[23] In contrast, another meta-analysis in advanced colorectal cancer has shown that PFS is an acceptable surrogate for survival in that disease, which underscores the fact that surrogacy is disease- and treatment-specific (Buyse et al, submitted for publication).

In addition to tumor response and PFS, we also investigated whether disease control or TTP could be used as surrogates for survival. These two end points are sometimes used to provide supporting evidence for regulatory approval. In our analyses, the results and conclusions for disease control and for TTP are similar to those obtained for tumor response and PFS, respectively.

The statistical methods adopted in the present paper and elsewhere use a correlation approach to the validation of potential surrogate end points.[23-24] As such, they offer quantitative evidence but do not result in any absolute claim of surrogacy.[20,25] Indeed, a potential limitation of using ρ is that there is no standard value above which a correlation can be claimed to be sufficient. It is helpful to consider the squared value of ρ, which is the variance in the variable explained by its predictor. For instance, the ρ between the treatment effects on response and on survival was 0.57; hence, less than one third ($0.57^2 = 0.32$) of the variability in treatment effects on survival can be explained by the variability in treatment effects on response. The wide confidence intervals around some ρ data appropriately reflect the substantial statistical uncertainty despite the number of trials and patients available for analysis. For instance, the 95% CI around 0.57 ranged from −0.31 to 1.44, which was too wide to be informative.

The present analyses showed that tumor response is an acceptable surrogate for PFS in patients with advanced breast cancer, as the association between these end points is quite strong, and that treatment effects on tumor response predict treatment effects on PFS extremely well (ρ, 0.96; CI 0.73 to 1.19). However, the clinical implications of this finding are rather limited. If surrogates are sought for survival, they will have to be validated by using data from other randomized trials.

AUTHOR CONTRIBUTIONS

AQ: F

Conception and design: Tomasz Burzykowski, Marc Buyse, Martine J. Piccart-Gebhart
Provision of study materials or patients: Martine J. Piccart-Gebhart, George Sledge, James Carmichael, Hans-Joachim Lück, John R. Mackey, Jean-Marc Nabholtz, Robert Paridaens, Laura Biganzoli, Jacek Jassem, Marijke Bontenbal, Jacques Bonneterre, Stephen Chan, Gul Atalay Basaran, Patrick Therasse
Collection and assembly of data: Tomasz Burzykowski
Data analysis and interpretation: Tomasz Burzykowski, Marc Buyse
Manuscript writing: Tomasz Burzykowski, Marc Buyse, Martine J. Piccart-Gebhart
Final approval of manuscript: Tomasz Burzykowski, Marc Buyse, Martine J. Piccart-Gebhart, George Sledge, James Carmichael, Hans-Joachim Lück, John R. Mackey, Jean-Marc Nabholtz, Robert Paridaens, Laura Biganzoli, Jacek Jassem, Marijke Bontenbal, Jacques Bonneterre, Stephen Chan, Gul Atalay Basaran, Patrick Therasse

| balt5/zlj-jco/zlj-jco/zlj01208/zlj6842-08a | kroenink | S=3 | 3/11/08 | 10:36 | 4/Color Figure(s): F1-4 | Art: 08407 | Input-NMV |

Burzykowski et al

REFERENCES

1. Carmichael J. UKCCCR trial of epirubicin and cyclophosphamide (EC) vs. epirubicin and Taxol (ET) in the first line treatment of women with metastatic breast cancer (MBC). J Clin Oncol 20:22a, 2001 (abstr 84)

2. Luck HJ, Thomssen C, Untch M, et al: Multicentric phase III study in first line treatment of advanced breast cancer (ABC): Epirubicin/Paclitaxel (ET) vs Epirubicin/Cyclophosphamide (EC)—A study of the AGO Breast Cancer Group. J Clin Oncol 19:73, 2000 (abstr 280)

3. Biganzoli L, Cufer T, Bruning P, et al: Doxorubicin and paclitaxel versus doxorubicin and cyclophosphamide as first-line chemotherapy in metastatic breast cancer: The European Organization for Research and Treatment of Cancer 10961 multicenter phase III trial. J Clin Oncol 20:3114-3121, 2002

4. Jassem J, Pienkowski T, Pluzanska A, et al: Doxorubicin and paclitaxel versus fluorouracil, doxorubicin, and cyclophosphamide as first-line therapy for women with metastatic breast cancer: Final results of a randomized phase III multicenter trial. J Clin Oncol 19:1707-1715, 2001

5. Mackey JR, Paterson A, Dirix L, et al: Final results of the phase III randomized trial comparing docetaxel (T), doxorubicin (A) and cyclophosphamide (C) to FAC as first line chemotherapy (CT) for patients (pts) with metastatic breast cancer (MBC). J Clin Oncol 21:35a, 2002 (abstr 137)

6. Nabholtz JM, Falkson G, Campos D, et al: Docetaxel and doxorubicin compared with doxorubicin and cyclophosphamide as first-line chemotherapy for metastatic breast cancer: Results of a randomized, multicenter, phase III trial. J Clin Oncol 21:968-975, 2003

7. Bontenbal M, Creemers GJ, Braun HJ, et al: Phase II to III study comparing doxorubicin and docetaxel with fluorouracil, doxorubicin, and cyclophosphamide as first-line chemotherapy in patients with metastatic breast cancer: Results of a Dutch community setting trial for the clinical trial group of the Comprehensive Cancer Centre. J Clin Oncol 23:7081-7088, 2005

8. Bonneterre J, Dieras V, Tubiana-Hulin M, et al: Phase II multicenter randomized study of deocetaxel plus epirubicin versus fluorouracil plus epirubicin and cyclophosphamide in metastatic breast cancer. Br J Cancer 91:1466-1471, 2004

9. Sledge GW, Neuberg D, Bernardo P, et al: Phase III trial of doxorubicin, paclitaxel and the combination of doxorubicin and paclitaxel as front-line chemotherapy for metastatic breast cancer: An intergroup trial (E1193). J Clin Oncol 21:588-592, 2003

10. Paridaens R, Biganzoli L, Bruning P, et al: Paclitaxel versus doxorubicin as first-line single-agent chemotherapy for metastatic breast cancer: A European Organization for Research and Treatment of Cancer randomized study with cross-over. J Clin Oncol 18:724-733, 2000

11. Chan S, Friedrichs K, Noel D, et al: Prospective randomized trial of docetaxel versus doxorubicin in patients with metastatic breast cancer. J Clin Oncol 17:2341-2354, 1999

12. WHO Handbook for Reporting Results of Cancer Treatment. Geneva, Switzerland, World Health Organization, WHO Offset Publication No. 48, 1979

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

14. Burzykowski T, Molenberghs G, Buyse M (eds): Evaluation of Surrogate Endpoints. New York, NY, Springer, 2005

15. Burzykowski T, Molenberghs G, Buyse M: The validation of surrogate end points using data from randomized clinical trials: A case-study in advanced colorectal cancer. J R Stat Soc Ser A Stat Soc 167:103-124, 2004

16. Buyse M, Piedbois P: On the relationship between response to treatment and survival time. Stat Med 15:2797-2812, 1996

17. Burzykowski T, Molenberghs G, Buyse M, et al: Validation of surrogate endpoints in multiple randomized clinical trials with failure-time end points. J R Stat Soc C 50:405-422, 2001

18. Di Leo A, Bleiberg H, Buyse M: Overall survival is not a realistic endpoint for clinical trials in advanced solid tumors: A critical assessment based on recently reported phase III trials in colorectal and breast cancer. J Clin Oncol 21:2045-2047, 2003

19. Hackshaw A, Knight A, Barrett-Lee P, et al: Surrogate markers and survival in women receiving first-line combination anthracycline chemotherapy for advanced breast cancer. Br J Cancer 93:1215-1221, 2005

20. Prentice RL: Surrogate end points in clinical trials: Definitions and operational criteria. Stat Med 8:431-440, 1989

21. Buyse M, Molenberghs G: Criteria for the validation of surrogate end points in randomized experiments. Biometrics 54:1014-1029, 1998

22. Bruzzi P, Del Mastro L, Sormani MP, et al: Objective response to chemotherapy as a potential surrogate end point of survival in metastatic breast cancer patients. J Clin Oncol 23:5117-5125, 2005

23. Buyse M, Thirion P, Carlson RW, et al: Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: A meta-analysis. Lancet 356:373-378, 2000

24. Sargent D, Wieand S, Haller DG, et al: Disease-free survival versus overall survival as a primary endpoint for adjuvant colon cancer studies: Individual patient data from 20,898 patients on 18 randomized trials. J Clin Oncol 23:8664-8670, 2005

25. Baker SG: Surrogate endpoints: Wishful thinking or reality? J Natl Cancer Inst 98:502-503, 2006   **AQ: G**
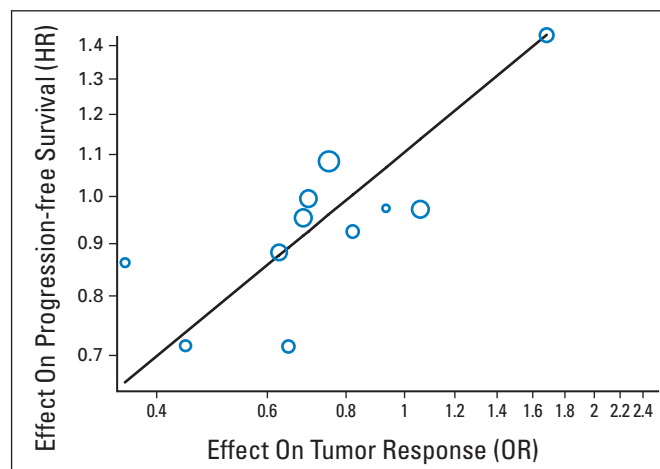
FA1

Appendix

Fig A1. Correlation between treatment effects on response and on PFS. Each circle represents a trial, and the surface area of the circle is proportional to the size of the corresponding trial.