Circulating Tumor Cells versus Imaging—Predicting Overall Survival in Metastatic Breast Cancer

G. Thomas Budd,[1] Massimo Cristofanilli,[2] Mathew J. Ellis,[3] Allison Stopeck,[4] Ernest Borden,[1] M. Craig Miller,[5] Jeri Matera,[5] Madeline Repollet,[5] Gerald V. Doyle,[5] Leon W.M.M. Terstappen,[5] and Daniel F. Hayes[6]

Abstract

Purpose: The presence of ≥5 circulating tumor cells (CTC) in 7.5 mL blood from patients with measurable metastatic breast cancer before and/or after initiation of therapy is associated with shorter progression-free and overall survival. In this report, we compared the use of CTCs to radiology for prediction of overall survival.

Experimental Design: One hundred thirty-eight metastatic breast cancer patients had imaging studies done before and a median of 10 weeks after the initiation of therapy. All scans were centrally reviewed by two independent radiologists using WHO criteria to determine radiologic response. CTC counts were determined ~4 weeks after initiation of therapy. Specimens were analyzed at one of seven laboratories and reviewed by a central laboratory.

Results: Interreader variability for radiologic responses and CTC counts were 15.2% and 0.7%, respectively. The median overall survival of 13 (9%) patients with radiologic nonprogression and ≥5 CTCs was significantly shorter than that of the 83 (60%) patients with radiologic nonprogression and <5 CTCs (15.3 versus 26.9 months; $P = 0.0389$). The median overall survival of the 20 (14%) patients with radiologic progression and <5 CTCs was significantly longer than the 22 (16%) patients with ≥5 CTCs that showed progression by radiology (19.9 versus 6.4 months; $P = 0.0039$).

Conclusions: Assessment of CTCs is an earlier, more reproducible indication of disease status than current imaging methods. CTCs may be a superior surrogate end point, as they are highly reproducible and correlate better with overall survival than do changes determined by traditional radiology.

Early detection and more effective local and systemic therapies have resulted in a declining death rate from breast cancer (1). Despite this improving prognosis, breast cancer is incurable when metastatic and is the second leading cause of cancer death among U.S. women (1). At present, treatment of patients with metastatic breast cancer is palliative, involving judiciously applied serial endocrine, chemotherapeutic, or biological therapies in an attempt to produce a series of remissions (2). Treatment with an individual agent or combination of agents is generally continued as long as the patient shows no evidence of disease progression or unacceptable toxicity. Because the treatment of breast cancer is associated with significant toxicity, it is important to be able to assess treatment efficacy in individual patients so that effective therapy can be continued and ineffective therapy discontinued. Similarly, when investigating novel treatments for breast cancer, the ability to assess the efficacy of therapy is essential to evaluate the effectiveness of the treatment under investigation. At present, either in routine clinical practice or in the research setting, treatment efficacy is assessed by serial history and physical examinations, radiographic studies, and evaluation of serologic "tumor markers" (2). Each of these approaches is confounded by inaccuracies. Simpler, more effective means of assessing the effectiveness of therapy are needed, both to allow optimal treatment of individual patients and to serve as surrogate measures of activity for novel therapies.

We have previously reported a novel, automated immunomagnetic technique to detect circulating tumor cells (CTC) in the blood of patients with breast and other cancers (3). In a prospective trial done in patients with measurable metastatic breast cancer, we found that the number of CTCs was a powerful prognostic indicator for such patients (4, 5). In the current article, we report (*a*) the reproducibility of each of these methods, (*b*) the correlation between the presence of CTCs and objective response to therapy as assessed using serial radiographic studies, and (*c*) a comparison of the correlation of each of these means of evaluating treatment success with overall survival.

Authors' Affiliations: [1]Cleveland Clinic Foundation, Cleveland, Ohio; [2]M.D. Anderson Cancer Center, Houston, Texas; [3]Washington University, St. Louis, Missouri; [4]University of Arizona, Tucson, Arizona; [5]Immunicon Corp., Huntingdon Valley, Pennsylvania; and [6]University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan

Received 8/12/05; revised 3/9/06; accepted 5/2/06.

Grant support: Immunicon, Inc. (Huntingdon Valley, PA).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: G. Thomas Budd, Cleveland Clinic Foundation, 9500 Euclid Avenue, Desk R35, Cleveland, OH 44195. Phone: 216-444-6480; Fax: 216-445-2369; E-mail: buddg@ccf.org.

doi:10.1158/1078-0432.CCR-05-1769

Materials and Methods

Study design. As previously reported, a prospective, longitudinal, double-blind, clinical trial was conducted at 20 academic and private centers throughout the United States (4, 5). This trial studied the use of CTCs to predict progression-free and overall survival in metastatic breast cancer patients starting a new line of therapy. One hundred thirty-eight of the 177 patients enrolled in the trial had imaging studies done at baseline and periodically throughout their treatment. Reasons for not having serial studies available for review included death before the first follow-up imaging (12 patients), scans read as indeterminant after central imaging review (9 patients), scans not provided for central imaging review (8 patients), patients being found to have bone-only disease (4 patients), patients being lost to follow-up before the first follow-up imaging (5 patients), and 1 patient whose treatment was changed before the first follow-up imaging.

Principal inclusion criteria were women ≥18 years of age, progressive measurable metastatic breast cancer, commencement of a new systemic therapy, and an Eastern Cooperative Oncology Group performance status score of 0 to 2. Prior adjuvant therapy, prior treatment for metastatic disease, or both were permitted, as radiographic studies are used to assess response following all lines of therapy and there was no *a priori* reason to assume that CTCs should be used differently. The institutional review board at each center approved the study protocol, and all patients provided written informed consent. All participating investigators were experienced in clinical research, and the study was prospectively monitored for compliance by an independent clinical research organization (Medical Device Consultants, Inc., North Attleboro, MA).

Before starting a new systemic treatment, all patients had computed tomography and/or magnetic resonance imaging scans of the chest and abdomen, a whole body bone scan, and a baseline blood draw for enumeration of CTCs. Reassessment of disease status by the same clinical and radiologic modalities used at baseline was conducted every 9 to 12 weeks depending on treatment type and schedule. Serial blood specimens were collected at approximately monthly intervals for a period of up to 6 months. Patients remained on study for 6 months or until they progressed or died. Clinicians were blinded to CTC results and to the independent assessment of radiographs, and all decisions about patient care were made using standard clinical and radiographic evaluation done at the clinical sites.

Measurement of patient responses by radiologic assessment. A core imaging facility (Beacon Bioscience, Inc., Doylestown, PA) audited all sites to ensure that imaging equipment conformed to Digital Imaging and Communications in Medicine standards and that all operators had the proper state certification. The protocol specified an overall blinded read design for independent tumor assessment from medical images acquired during the trial. This stipulation included operator training and a centralized review of imaging studies at the core imaging facility by two independent radiologists not involved with the clinical aspects of the trial. A third independent radiologist adjudicated disagreements between the two primary readers about progressive disease and nonprogressive disease. Readers were blinded to all patient, clinical, and CTC information. All reads were digital, with readers informed of the chronology of the image. Readers identified up to eight lesions per patient per time point by describing the longest dimension of the lesion and the longest perpendicular dimension. These two dimensions were multiplied, and the "cross-product" was reported. Summed measurements for the cross-products were calculated, and percentage change from the previous time point was determined. Overall response was determined by the independent readers according to the modified WHO response criteria (6). Although all patients had measurable disease, nonmeasurable lesions were included in the determination of patient status as described in the WHO guidelines. Progressive disease was defined as a ≥25% increase in the sum of all lesions or appearance of a new measurable or nonmeasurable lesion. Partial response was defined as a decrease in the sum of all lesions of ≥50% and no new lesions. Image presentation and analysis were done using Cheshire software (Parexel, Inc., Waltham, MA). All data were collected via an electronic case report form, which was also used to control image displays to guarantee that they were correctly displayed. The database architecture ensured that the readers answered the appropriate questions and autochecked for missing data. After committing the data to the database, the random code number was removed from the queue, the entry panel was locked, and changes could no longer be made to the evaluation. All reads were conducted at the core facility and monitored by imaging specialists.

Isolation and enumeration of CTC. As previously described, blood samples were drawn into 10 mL EDTA Vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ), to which a cellular preservative was subsequently added (3, 4). Samples were maintained at room temperature and processed within 72 hours after collection. All CTC evaluations were done in one of two central laboratories (Immunicon Corp., Huntingdon Valley, PA or IMPATH Predictive Oncology, Los Angeles, CA) or at one of five participating cancer centers (M.D. Anderson Cancer Center, Houston, TX; Taussig Cancer Center, Cleveland Clinic, Cleveland, OH; Duke University Cancer Center, Charlotte, NC; University of Arizona, Tucson, AZ; or University of Michigan, Ann Arbor, MI). The CellSearch System (Veridex LLC, Raritan, NJ) was used for the isolation and enumeration of CTC. The system consisted of a semiautomated sample preparation system and the CellSearch Epithelial Cell kit to immunomagnetically enrich cells expressing the epithelial cell adhesion molecule (7). Isolated cells were then fluorescently labeled with the nucleic acid dye 4′,6-diamidino-2-phenylindole and labeled monoclonal antibodies specific for leukocytes (CD45-allophycocyan) and epithelial cells (cytokeratin 8,18,19-phycoerythrin). Identification and enumeration of CTCs was done using the CellSpotter Analyzer (Immunicon), a semiautomated fluorescence microscopy system that permits computer-generated reconstruction of cellular images (8, 9). CTCs are defined as nucleated cells lacking CD45 and expressing cytokeratin and were enumerated by trained operators blinded to patient outcomes. Operational details of the CellSearch and CellSpotter systems and preclinical performance have been described previously (3, 4, 7–9). Before initiation of the study, operators at the testing sites were trained by Immunicon on the classification of images generated by the CellSpotter Analyzer. Training was continued throughout the study by sending sets of images on a monthly basis to the testing sites for review and performance assessment.

For CTCs, a threshold of ≥5 CTCs/7.5 mL blood, the derivation of which has been previously described in detail (4), was used for evaluation of the results. For each blood sample tested, digital images produced by the CellSpotter Analyzer were reviewed by an operator at each testing site and CTCs were tabulated. All images from each individual experiment were copied to a single compact disc, which was then sent to a central laboratory and reviewed by an operator blinded to clinical and CTC results from the testing site.

Analysis of interreader and intrareader variability. Independent radiologic review classified each patient according to WHO criteria as having had progressive disease, stable disease, complete response, partial response, or indeterminate at the first radiographic follow-up time point. No complete responses were observed so that only partial responses will be referred to hereafter. Twenty-five patients (roughly equally divided between progressors and nonprogressors) were randomly selected for evaluation of interobserver and intraobserver variability. For these 25 patients, each radiologist determined the response at three separate sittings, each sitting separated by a minimum of 1 week. The radiologic responses at each sitting for each radiologist were compared as indeterminate versus stable disease/partial response versus progressive disease. Interreader variability was determined by comparing the radiologic responses of the two radiologists, classified as indeterminate versus stable disease/partial response (nonprogression) versus progressive disease (progression). Intrareader variability was calculated by comparing the radiologic responses of the two radiologists in the subset of 25 patients.

Seventy-one of the 138 patients with evaluable imaging results had duplicate tubes drawn for CTC evaluation at the first follow-up blood draw. Interreader variability for the CTC counts at the first follow-up

blood draw was determined by counting the number of instances in which the operator at the testing site was not in concordance with the central laboratory for classification of a sample as progression (≥5 CTCs) versus nonprogression (<5 CTCs). Tube-to-tube variation in the same patient was determined in a subset of 71 patients by drawing two tubes of blood at the same time. The specimens were then processed and classified as CTC progression versus CTC nonprogression in each of the two tubes by both the site and the central laboratory. The results were then compared.

Statistical analysis. Overall survival was defined as the elapsed time between the date of the baseline blood draw and either the date of death or the last follow-up (if death was not observed during the follow-up period). Survival curves were compared using log-rank testing. The median overall survival (or 50th percentile of the survival time) was determined from the Kaplan-Meier product limit estimate of the survivor functions. Cox proportional hazards regression analysis was used to determine hazard ratios for overall survival. The distribution of patients above and below the CTC threshold who died within 6 and/or 12 months from the time of the baseline blood draw was compared using Fisher's exact test. All *P*s reported are two sided, and all averages are presented as the mean ± SD.

Results

Patient characteristics. Patient demographics are shown in Table 1. The average time on study for each patient was 4.1 ± 1.8 months, and at the time of this analysis, 73 (53%) of the patients had died. The average time between the baseline blood draw and the follow-up radiologic disease reevaluation was 11.9 ± 5.7 weeks (median, 10.1; range, 1.9-34.1). At that time point, the independent radiologic review scored 26 (19%) patients as having a partial response, 70 (51%) as having stable disease, and 42 (30%) as having progressive disease. The average time between the baseline blood draw and the first follow-up blood draw was 4.6 ± 2.5 weeks (median, 4.0; range, 1.4-16.9). At baseline, 60 (43%) patients had ≥5 CTCs/7.5 mL blood, whereas at the first follow-up blood draw, 35 (25%) of the patients had ≥5 CTCs/7.5 mL blood.

Table 1. Patient demographics (*N* = 138)

Age, y (median)	58.5 + 13.6 (59)
Race, *n* (%)	
White	118 (86)
Black	13 (9)
Hispanic	6 (4)
Unknown	1 (1)
Hormone receptor, *n* (%)	
+	94 (68)
−	44 (32)
Her2, *n* (%)	
Her2− (0, 1+)	81 (59)
Her2+ (2+, 3+)	34 (24)
Unknown	23 (17)
Type of therapy, *n* (%)	
Hormone/immunotherapy	46 (33)
Chemotherapy alone/combination	92 (67)
Line of therapy, *n* (%)	
1st line	67 (49)
≥2nd line	71 (51)
Site metastasis, *n* (%)	
Visceral	122 (88)
Nonvisceral	16 (12)
Status at last follow-up, *n* (%)	
Alive	65 (47)
Dead	73 (53)
Average follow-up time (mo)	
Alive	19.8 ± 5.4 (range, 2.7-28.8)
Dead	12.2 ± 6.5 (range, 1.9-26.9)

Interreader and intrareader variability of radiologic responses and CTC counts. Table 2 summarizes the interreader and intrareader discordance between the two radiologists as well as the interreader and tube-to-tube variability for the CTC assay. To determine if the interreader variability of radiologic and CTC assessment as measured was truly representative, we included imaging segments from end of treatment assessments in all 138 patients. Additional CTC assessments from pretreatment and post-first follow-up visits were also analyzed, and the data are presented in Table 2.

Relation between tumor load measured by bidimensional radiography and CTCs. The bidimensional sum of the metastatic lesions at the first follow-up disease assessment in the 138 cases was measured by each of the two radiologists. Tumor size measured by radiologist 1 ranged from 49 to 19,067 mm^2 (mean, 2,383 ± 3,258) and by radiologist 2 from 20 to 21,969 mm^2 (mean, 2,027 ± 3,505). A reasonable correlation was found between the measurements of both radiologists (r^2 = 0.78, slope of 0.94). CTC levels per 7.5 mL blood in these patients ranged from 0 to 9,864 at the first follow-up blood draw (mean, 104 ± 846; median, 0). No correlation existed between radiographic measurement of tumor load and CTC levels [radiologist 1 versus CTC: r^2 = 0.01, slope of 0.42, Spearman's rank correlation coefficient = 0.1461 (P = 0.1098); radiologist 2 versus CTC: r^2 = 0.001, slope of −0.17, Spearman's rank correlation coefficient = 0.0157 (P = 0.8661)].

Relation between therapy response measured by radiology and CTCs. Independent review of radiologic response was concordant with CTC levels in 105 of 138 (76%) cases. Eighty-three (60%) cases were found to have stable disease/partial response by radiologic criteria and <5 CTCs/7.5 mL blood, and 22 (16%) cases had progressive disease by radiographic criteria and ≥5 CTCs/7.5 mL blood. Of the 33 (24%) discrepant cases, 20 (14%) with progressive disease by radiographic criteria had <5 CTCs/7.5 mL blood and 13 (9%) with stable disease/partial response by radiographic criteria had ≥5 CTCs/7.5 mL blood.

Comparison of CTCs and radiology to predict survival. Table 3 summarizes the survival of the patients according to their radiographic response and CTC count at the time of first follow-up. Figure 1 shows the Kaplan-Meier curves for overall survival in the 138 patients according to their CTC count at the first follow-up visit after initiation of therapy (Fig. 1A) and radiologic response at the follow-up disease reassessment after initiation of therapy (Fig. 1B). The median overall survival of 22.6 months [95% confidence interval (95% CI), 20.1 to >25.0] for the 103 (75%) patients with <5 CTCs was similar to the median overall survival of 24.9 months (95% CI, 20.1 to >25.0) for the 96 (70%) patients with a radiologic response of stable disease/partial response. In contrast, the median overall survival for the 35 (25%) patients with ≥5 CTCs was only 8.5 months (95% CI, 6.2-15.1) and was shorter than the median survival of 12.9 months (95% CI, 7.1-19.3) for the 42 (30%) patients with a radiologic response of progressive disease.

To compare the predictive values for overall survival provided by CTC versus radiographic criteria, we used Cox proportional hazards regression. For radiology, the univariate Cox hazard ratio for death in patients with progressive disease was 2.59 (95% CI, 1.63-4.14; χ^2 = 14.95; P = 0.0001) compared

Table 2. Variability of radiologic responses and CTC enumeration

Radiology	Imaging segment(s)	*n*	% Disagreement (I vs S/PR vs PD)
Interreader			
Reader 1 vs reader 2	1st follow-up only	138	15.2
	All follow-ups	235	17.0
Intrareader			
Reader 1	1st follow-up	25	24.0
	All follow-ups	32	21.9
Reader 2	1st follow-up only	25	20.0
	All follow-ups	32	21.9
CTCs	**Blood draw(s)**	***n***	**% Disagreement (<5 CTCs vs ≥5 CTCs)**
Interreader			
Site vs central (tube 1)	1st follow-up only	138	0.7
	All follow-ups	695	1.0
Site vs central (tube 2)	1st follow-up only	71	0.0
	All follow-ups	409	1.0
Tube 1 vs tube 2			
Site	1st follow-up only	71	5.6
	All follow-ups	403	5.5
Central	1st follow-up only	71	4.2
	All follow-ups	403	5.5

Abbreviations: I, indeterminate; S/PR, stable disease/partial response; PD, progressive disease.

with a univariate Cox hazard ratio for death in patients with ≥5 CTCs of 3.18 (95% CI, 1.96-5.96; χ^2 = 19.26; P < 0.0001). When CTCs and radiology were combined in a multivariate Cox proportional hazards regression, the χ^2 for the model was 26.51 (P < 0.0001), with a hazard ratio of 2.00 (95% CI, 1.22-3.29; P = 0.006) for the radiology assessment and 2.53 (95% CI, 1.51-4.22; P < 0.001) for CTCs.

To determine whether the line of therapy influenced the ability of CTCs and/or radiology to predict overall survival, Kaplan-Meier curves were generated for both groups and are presented in Fig. 2A-D. In the 67 (49%) patients starting their first line of therapy for metastatic disease (endocrine or chemotherapy), the median overall survival of >25.0 months (95% CI, 21.9 to >25.0) for the 52 (78%) patients with <5 CTCs (Fig. 2A) was similar to the median overall survival of 24.9 months (95% CI, 21.9 to >25.0) for the 54 (81%) patients with a radiologic response of stable disease/partial response (Fig. 2B). Likewise, the median overall survival of 12.9 months (95% CI, 6.2-19.5) for the 15 (22%) patients with ≥5 CTCs (Fig. 2A) was similar to the median overall survival of 12.9 months (95% CI, 5.8-19.3) for the 13 (19%) patients with a radiologic response of progressive disease (Fig. 2B). The difference between the median overall survivals of the two patient groups was significant for both CTCs and radiologic response (Ps ≤ 0.0001, log-rank test).

In the 72 (51%) patients starting their second or higher line of therapy, the median overall survival of 19.9 months (95% CI, 14.3-22.6) for the 51 (71%) patients with <5 CTCs (Fig. 2C) was similar to the median overall survival of 18.0 months (95% CI, 13.9-26.9) for the 42 (59%) patients with a radiologic response of stable disease/partial response (Fig. 2D). In contrast, the median overall survival of 6.4 months (95% CI, 3.3-10.9) for the 20 (29%) patients with ≥5 CTCs (Fig. 2C) was much shorter than the median overall survival of 13.4 months (95% CI, 6.4-21.6) for the 29 (41%) patients with a radiologic response of progressive disease (Fig. 2D). The difference

Table 3. Radiologic response versus CTC enumeration to predict overall survival in 138 patients with measurable metastatic breast cancer

Radiographic response	*n* (%)	Mortality at 6 months (%)*	Mortality at 12 months (%)†	Median overall survival, mo (95% CI)
S/PR	96 (70)	5/95 (5)	18/90 (20)	24.9 (20.1 to >25.0)
PD	42 (30)	11/41 (26)	19/41 (46)	12.9 (7.1-19.3)
Fisher's exact *P*		0.001	0.003	*P* < 0.0001, log-rank test
CTC levels/7.5 mL at 1st follow-up	***n* (%)**	**Mortality at 6 months (%)***	**Mortality at 12 months (%)†**	**Median overall survival, mo (95% CI)**
<5 CTC	103 (75)	5/101 (5)	16/97 (16)	22.6 (20.1 to >25.0)
≥5 CTC	35 (25)	11/35 (31)	21/34 (62)	8.5 (6.2-15.1)
Fisher's exact *P*		<0.001	<0.001	*P* < 0.0001, log-rank test

*Two of the patients who were last known to be alive had <6 months of total follow-up time and were excluded from this analysis.
†Seven of the patients who were last known to be alive had <12 months of total follow-up time and were excluded from this analysis.

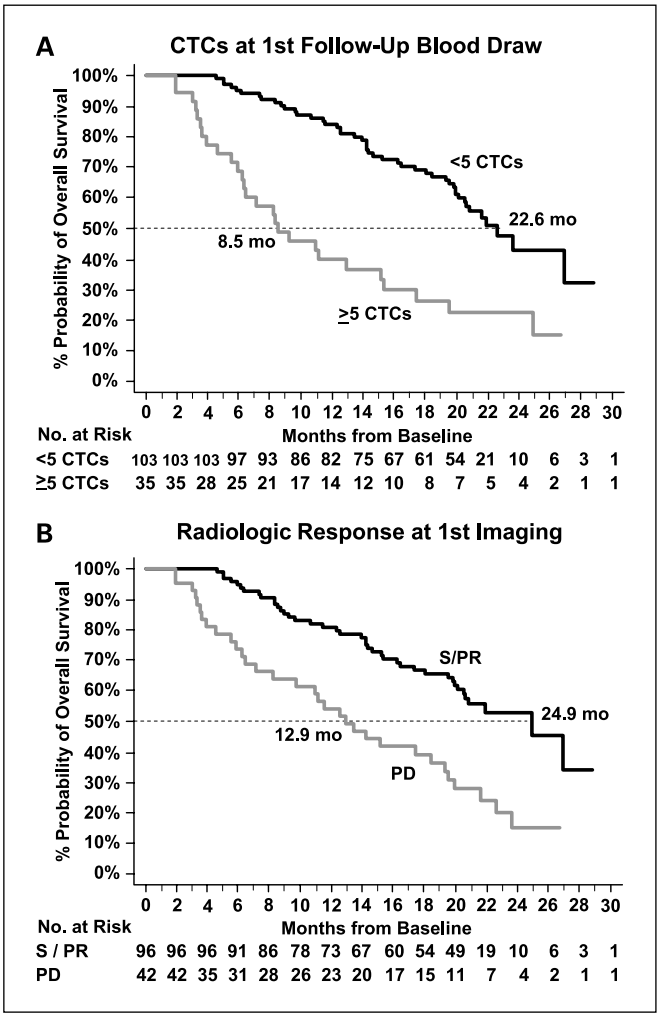

Fig. 1. Overall survival time was calculated from the date of the baseline blood draw taken before the initiation of a new line of therapy. *A,* probability of overall survival for patients with <5 CTCs versus ≥5 CTCs per 7.5 mL blood at the first follow-up blood draw ($P < 0.0001$, log-rank test; Cox hazard ratio = 3.18; $\chi^2 = 19.26$; $P < 0.0001$). *B,* probability of overall survival for patients with stable disease/partial response (*S/PR*) versus progressive disease (*PD*) by radiologic imaging at the first disease reassessment ($P < 0.0001$, log-rank test; Cox hazard ratio = 2.59; $\chi^2 = 14.95$; $P = 0.0001$).

between the median overall survivals of the two patient groups was significant for CTCs ($P = 0.0009$, log-rank test) but not for radiologic response ($P = 0.2209$, log-rank test).

Table 4 shows the median overall survivals as well as the percentages of patients who died within 6 and 12 months from

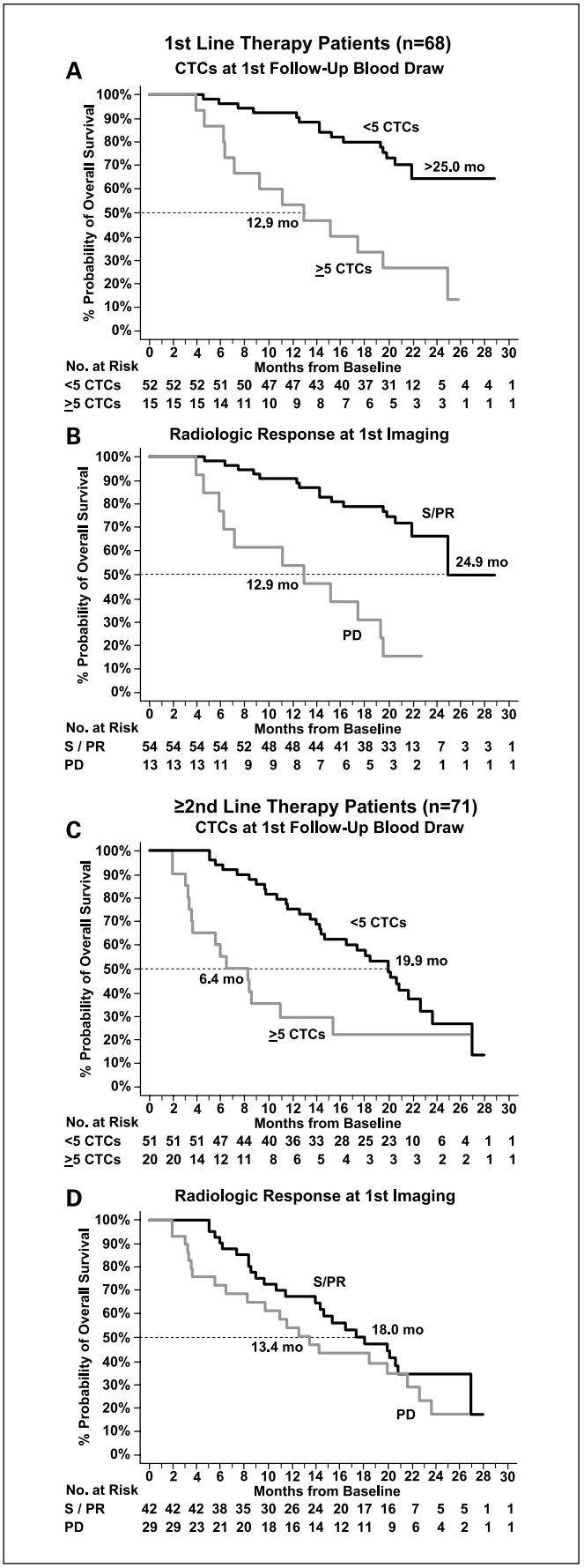

Fig. 2. Overall survival time was calculated from the date of the baseline blood draw taken before the initiation of a new line of therapy. *A,* probability of overall survival for 67 patients on their first line of therapy with <5 CTCs versus ≥5 CTCs per 7.5 mL blood at the first follow-up blood draw ($P = 0.0001$, log-rank test; Cox hazard ratio = 4.08; $\chi^2 = 11.30$; $P = 0.0008$). *B,* probability of overall survival for 67 patients on their first line of therapy with stable disease/partial response versus progressive disease by radiologic imaging at the first disease reassessment ($P < 0.0001$, log-rank test; Cox hazard ratio = 5.37; $\chi^2 = 14.71$; $P = 0.0001$). *C,* probability of overall survival for 71 patients on their second or higher line of therapy with <5 CTCs versus ≥5 CTCs per 7.5 mL blood at the first follow-up blood draw ($P = 0.0009$, log-rank test; Cox hazard ratio = 2.79; $\chi^2 = 8.86$; $P = 0.0029$). *D,* probability of overall survival for 71 patients on their second or higher line of therapy with stable disease/partial response versus progressive disease by radiologic imaging at the first disease reassessment ($P = 0.2209$, log-rank test; Cox hazard ratio = 1.44; $\chi^2 = 1.46$; $P = 0.2277$).

Table 4. Radiologic response and CTC enumeration combined to predict overall survival in 138 patients with measurable metastatic breast cancer

Group	Radiographic response	CTC number at 1st follow-up	*n* (%)	Mortality at 6 months (%)*	Mortality at 12 months (%)†	Median overall survival, mo (95% CI)
1	S/PR	<5 CTC	83 (60)	3/82 (4)	12/78 (15)	26.9 (20.5 to >25.0)
2	PD	<5 CTC	20 (14)	2/19 (10)	4/19 (21)	19.9 (12.5-23.6)
3	S/PR	≥5 CTC	13 (9)	2/13 (15)	6/12 (50)	15.3 (6.3 to >25.0)
4	PD	≥5 CTC	22 (16)	9/22 (41)	15/22 (68)	6.4 (3.5-12.9)
—	Fisher's exact *P*			<0.001	<0.001	*P* < 0.0001, log-rank test

*Two of the patients who were last known to be alive had <6 months of total follow-up time and were excluded from this analysis.
†Seven of the patients who were last known to be alive had <12 months of total follow-up time and were excluded from this analysis.

the time of the baseline blood draw when the CTC counts at the first follow-up blood draw and the radiologic response at the first follow-up disease reassessment are combined. Group 1 consisted of 83 (59%) patients with <5 CTCs and a radiologic response of stable disease/partial response, group 2 consisted of 20 (14%) patients with <5 CTCs and a radiologic response of progressive disease, group 3 consisted of 13 (9%) patients with ≥5 CTCs and a radiologic response of stable disease/partial response, and group 4 consisted of 22 (16%) patients with ≥5 CTCs and a radiologic response of progressive disease. The differences in the percentages of patients who died within 6 and 12 months were highly significant ($P < 0.001$, Fisher's exact test). The difference in the median overall survivals between the four groups of patients was statistically significant ($P < 0.0001$, log-rank test).

A Kaplan-Meier plot comparing groups 1, 2, 3, and 4 is presented in Fig. 3. Within the group of "responding" patients by the CTC assay (i.e., <5 CTCs), radiographic responders (stable disease/partial response) and nonresponders (progressive disease) do not have a significantly different prognosis (group 1 versus group 2; $P = 0.0785$, log-rank test). Similarly, within the group of "nonresponding" patients by the CTC assay (i.e., ≥5 CTCs), radiographic responders (stable disease/partial response) and nonresponders (progressive disease) also do not have a significantly different prognosis (group 3 versus group 4; $P = 0.0777$, log-rank test). In contrast, within the group of "responding" patients by radiology (i.e., stable disease/partial response), CTC responders (<5 CTCs) and CTC nonresponders (≥5 CTCs) do have a significantly different prognosis (group 1 versus group 3; $P = 0.0389$, log-rank test). Similarly, within the group of "nonresponding" patients by radiology (i.e., progressive disease), CTC responders (<5 CTCs) and CTC nonresponders (≥5 CTCs) also have a significantly different prognosis (group 2 versus group 4; $P = 0.0039$, log-rank test).

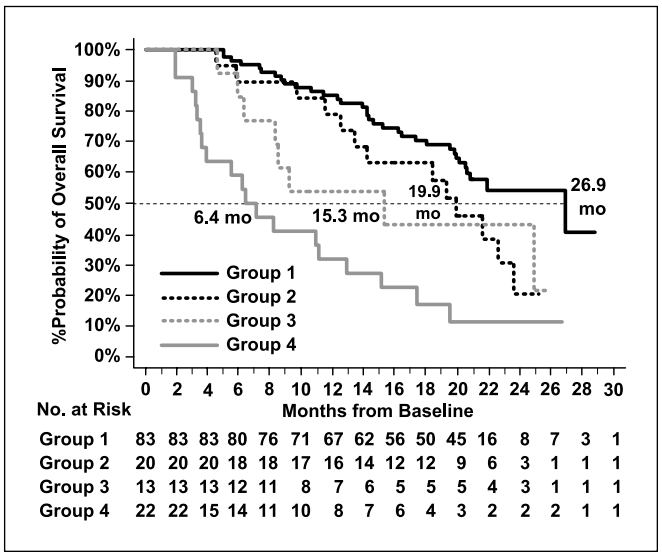

Fig. 3. Overall survival time was calculated from the date of the baseline blood draw taken before the initiation of a new line of therapy. Probability of overall survival for 83 (60%) patients with <5 CTCs and stable disease/partial response by radiology (*solid black line*, group 1), 20 (14%) patients with <5 CTCs and progressive disease by radiology (*dashed black line*, group 2), 13 (9%) patients with ≥5 CTCs and stable disease/partial response by radiology (*dashed gray line*, group 3), and 22 (16%) patients with ≥5 CTCs and progressive disease by radiology (*solid gray line*, group 4). Log-rank *P*s for a comparison of the survival curves are as follows: group 1 versus group 2 ($P = 0.0785$), group 1 versus group 3 ($P = 0.0389$), group 1 versus group 4 ($P < 0.0001$), group 2 versus group 3 ($P = 0.7549$), group 2 versus group 4 ($P = 0.0039$), and group 3 versus group 4 ($P = 0.0777$).

Discussion

The results reported here indicate that the evaluation of CTCs is an accurate measure of treatment efficacy. The CTC assay is both a more reproducible end point and a more robust surrogate of survival than is radiographic response.

Our data are consistent with data of others that indicate that radiologic response is subject to considerable technical limitations, such as intraobserver and interobserver variation (10, 11). When assay and imaging reproducibility are compared, CTCs showed a significant advantage with an interreader variability of 15% with imaging versus 1% with CTCs (Table 1). Accepting the validity of the CTC assay as a measure of treatment efficacy, this advantage in reproducibility implies that inaccurate disease status evaluations and inappropriate treatment decisions are less likely with the use of the CTC assay than with radiographic studies.

Traditionally, imaging has been used to measure the effectiveness of treatment in patients with metastatic breast cancer. Recognizing that radiographic disease stability may reflect either a beneficial effect of therapy or the presence of indolent disease, in this study we chose to define a radiologic response as the achievement of stable disease or a complete or partial response. This decision was based on (*a*) the recognition that the survival of patients with breast cancer with stable disease is equivalent to that of patients with radiographic tumor regression (12, 13) and (*b*) the clinical practice of continuing therapy as long as toxicity is acceptable and there is no evidence of disease progression (2, 12).

Tumor response status is a surrogate marker for a clinically meaningful benefit from treatment, such as quality of life or

survival (14–16). The relationship between any proposed surrogate and these clinically meaningful end points must be validated if the surrogate is to be used to assess or compare therapies. Our results indicate that serial assessment of CTCs in patients with metastatic breast cancer correlated with survival at least as well as serial radiographic assessment. Furthermore, the evaluation of CTCs provided prognostically significant information in patients who were felt to be stable or responding as well as those progressing by radiographic criteria (Fig. 3).

Our data suggest that CTC numbers reflect important aspects of the biology of an individual tumor. Prior studies have suggested that levels of classic circulating tumor markers, such as CA 15-3 or CA27.29, are related to tumor burden (17). However, our prior analyses failed to show a correlation between CTC and the number of sites of disease. Therefore, we conducted a more detailed analysis of tumor burden by summing the measurements of malignant lesions identified on radiographs by the independent reviewers at the restaging time point. Unlike soluble circulating tumor markers, such as assays that detect muc1, the number of CTCs is not simply a reflection of tumor bulk.

There are well-known difficulties associated with attempts to correlate radiographic response to therapy and survival (18, 19), and these difficulties could also apply to our analyses. Although it has often been found that patients showing radiographic improvement in association with a given treatment survive longer than those patients who fail to respond to treatment, there are fallacies that can arise if one concludes that the prolonged survival is a result of the response to treatment. First, both response to therapy and survival may be associated with common favorable prognostic factors, such as patient performance status and extent of prior therapy, so that the association of response and survival may simply represent a mutual correlation rather than a causal link. Second, because patients who die before the assessment of response will be considered nonresponders, there is an inherent bias favoring a correlation between response and longer survival. In our study, the number of CTCs was the most important prognostic factor to be associated with survival (4, 5), and we evaluated only those patients reaching the first assessment of response, minimizing the second problem described above. Although inherent problems in correlating response and survival remain, they are minimized by our analysis. We believe that response as assessed by enumeration of CTCs is clinically useful and seems to be as valid an assessment of treatment efficacy as is serial radiographic evaluation and assessment according to WHO criteria.

Our results have implications for both standard care and clinical research. More accurate determination of treatment effectiveness early in the course of therapy might spare patient toxicity from futile therapy and allow treatment to be changed to a more effective regimen. Whether such an early assessment of response results in an improved overall outcome or quality of life will need to be prospectively assessed in clinical trials designed to investigate this question. Moreover, all patients in this trial had measurable disease. We have just completed a prospective follow-up study of patients with nonmeasurable disease (20), who are notoriously difficult to follow (21). In these patients, a more reliable method to determine disease status would be of great value in helping the clinician decide whether to continue the present therapy or to change to an alternate treatment strategy. In addition to routine clinical use, CTCs might also be valuable in rapid response assessment in clinical investigations of novel agents. Indeed, one might even speculate that CTCs could become a validated end point for prospective clinical trials to be used instead of or to complement traditional end points, such as radiographic response, progression-free survival, or overall survival.

In conclusion, the number of CTCs in patients with metastatic breast cancer seems to reflect a disease state of biological and clinical importance. The assessment of CTCs has several advantages over serial radiographic evaluation. The CTC assay is more reproducible than is radiographic evaluation, shows useful results at an earlier time point than do radiologic studies, and seems to be a more robust predictor of survival than is radiographic response. The ability to serially quantitate and interrogate CTCs in patients with breast cancer makes possible new ways of managing and investigating this disease.

References

1. Jemal A, Murray T, Ward E, et al. Cancer statistics, 2005. CA Cancer J Clin 2005;55:10–30.
2. Ellis MJ, Hayes DF, Lippman ME. Treatment of metastatic breast cancer. In: Harris JR, Lippman ME, Morrow M, Osborne CK, editors. Diseases of the breast. 2nd ed. Philadelphia (PA): Lippincott; 1999. p. 749–97.
3. Allard WJ, Matera J, Miller MC. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with non-malignant diseases. Clin Cancer Res 2004;10:6897–904.
4. Cristofanilli M, Budd GT, Ellis M, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med 2004;351:781–91.
5. Cristofanilli M, Hayes DF, Budd GT, et al. Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. J Clin Oncol 2005; 23:1420–30.
6. Miller AB, Hoogstraten G, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981; 47:207–14.
7. Rao CG, Chianese D, Doyle GV, et al. Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors. Int J Oncol 2005;27:49–57.
8. Kagan M, Howard D, Bendele T, et al. A sample preparation and analysis system for identification of circulating tumor cells. J Clin Lig Assay 2002;25: 104–10.
9. Kagan M, Howard D, Bendele T, Rao C, Terstappen LWMM. Circulating tumor cells as cancer markers, a sample preparation and analysis system. In: Diamandis EP, Fritsche HA, Lilja H, Chan DW, Schwartz M, editors. Tumor markers: physiology, pathobiology, technology, and clinical applications. Washington (DC): AACC Press; 2002. p. 495–8.
10. Thiesse P, Ollivier L, Di Stefano-Louineau D, et al. Response rate accuracy in oncology trials: reasons for interobserver variability. Groupe Francais d'Immunotherapie of the Federation Nationale des Centres de Lutte Contre le Cancer. J Clin Oncol 1997;15:3507–14.
11. Erasmus JJ, Gladish GW, Broemeling L, et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. J Clin Oncol 2003; 21:2574–82.
12. Paterson AH, Cyr M, Szafran O, Lees AW, Hanson J. Response to treatment and its influence on survival in metastatic breast cancer. Am J Clin Oncol 1985;8:283–92.
13. Robertson J. Prognostic and response markers in the management of breast cancer. Cancer Treat Rev 1997;23 Suppl 1:S41–8.
14. Outcomes Working Group, Health Services Research Committee, American Society of Clinical Oncology. Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. J Clin Oncol 1996;14:671–9.
15. Nottage M, Siu LL. Principles of clinical trial design. J Clin Oncol 2002;20:42–6S.
16. Pazdur R. Response rates, survival, and chemotherapy trials. J Natl Cancer Inst 2000;92:1552–3.
17. Colomer R, Ruibal A, Salvador L. Circulating tumor marker levels in advanced breast carcinoma correlate with the extent of metastatic disease. Cancer 1989; 64:1674–81.
18. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol 1983;1:710–9.
19. Anderson JR. Commonly misused approaches to the analysis of cancer clinical trials. In: Crowley J, editor. Handbook of statistics in clinical oncology. New York: Marcel Dekker, Inc.; 2001. p. 491–502.
20. Budd GT, Cristofanilli M, Ellis MJ, et al. Monitoring circulating tumor cells (CTC) in non-measurable metastatic breast cancer (MBC). Proc Annu Meet Am Soc Clin Oncol 2005;24:A503.
21. Simon R. Clinical trials in cancer. In: Devita VT, Jr., Hellman S, Rosenberg SA, editors. Cancer: principles and practice of oncology. 6th ed. Philadelphia: Lippincott Williams and Wilkins; 2001. p. 521–38.

Clinical Cancer Research

Circulating Tumor Cells versus Imaging––Predicting Overall Survival in Metastatic Breast Cancer

G. Thomas Budd, Massimo Cristofanilli, Mathew J. Ellis, et al.

Clin Cancer Res 2006;12:6403-6409.