CORRESPONDENCE

Clinical Trial Designs for Cytostatic Agents

To the Editor: We read with great interest the special article on clinical trial designs for cytostatic agents written by Cancer Therapy Evaluation Program (CTEP) investigators,[1] having recently authored a similar article ourselves.[2] It is widely recognized that the historical approach to oncology drug development cannot be explicitly used for putative cytostatic agents. This is particularly applicable to phase II trials, where the historical approach of using small cohorts to look for a minimum response rate will clearly fail for drugs for which a classical partial response is not anticipated.[3] Korn et al[1] suggest that progression-free survival in a 30- to 50-patient cohort can be compared with that of historical controls. In our opinion, this approach is unlikely to be of benefit because of selection bias, lack of standardized data collection in many historical series, as well as the recent change in criteria for response and progression developed in part by CTEP investigators.[4]

A second approach suggested by the CTEP investigators is the use of small randomized screening studies with large type I errors (alpha = 0.2). This approach will result in a minimum positive rate of 20%. Phase III trials will then be required to sort out true positives from false positives. Historically, far less than 20% of drugs tested in the phase II setting were subsequently proven to be effective. Because phase III trials are notoriously expensive in terms of both financial and patient resources and because a large number of putative cytostatic agents are currently in development, it is unlikely that the current oncologic clinical research environment could support this number of phase III trials.

We were particularly concerned by the CTEP investigators' general recommendation against the randomized discontinuation trial design,[5] which has been extensively used outside of oncology, including the Prospective Randomized Study of Ventricular Function and Efficacy of Digoxin trial, demonstrating the efficacy of digoxin in chronic congestive heart failure.[6] This design is currently being evaluated in a Cancer and Leukemia Group B trial (sponsored by CTEP) in metastatic renal cancer. Accrual has been brisk, and the study has been well accepted by patients, in contrast to a National Cancer Institute trial of bevacizumab in a similar patient population, where patients are randomized to placebo verses active treatment.[7] Korn et al[1] are concerned that this design "will lead to an effective agent being declared ineffective if its continued use is not sufficiently better than its initial use." We would argue that a cytostatic agent that is only effective during a very short initial exposure period is highly unlikely to have a significant effect on disease progression or patient survival (as determined in a more standard phase III trial).

Given the paucity of success to date in the development of these promising agents, it will be crucial to evaluate a number of different clinical trial designs in a prospective and rigorous manner. Success of a phase II design will depend not only on its ability to stand up to critical statistical analysis but also on its ability to accrue rapidly and on its ability to accurately predict drug benefit in phase III trials. We strongly encourage an open-minded approach that places as much value on ingenuity and originality in trial design as is currently placed on target identification and validation.

Mark J. Ratain
Walter M. Stadler
University of Chicago
Chicago, IL

REFERENCES

1. Korn EL, Arbuck SG, Pluda JM, et al: Clinical trial designs for cytostatic agents: Are new approaches needed? J Clin Oncol 19:265-272, 2001
2. Stadler WM, Ratain MJ: Development of target-based antineoplastic agents. Invest New Drugs 18:7-16, 2000
3. Ratain MJ, Mick R, Schilsky RL, et al: Statistical and ethical issues in the design and conduct of phase I and II clinical trials of new anticancer agents. J Natl Cancer Inst 85:1637-1643, 1993
4. Therasse P, Arbuck SG, Eisenhauer EA, et al: New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 92:205-216, 2000
5. Kopec JA, Abrahamowicz M, Esdaile JM: Randomized discontinuous trials: Utility and efficiency. J Clin Epidemiol 46:959-971, 1993
6. Uretsky BF, Young JB, Shahidi FE, et al: Randomized study assessing the effect of digoxin withdrawal in patients with mild to moderate chronic congestive heart failure: Results of the PROVED trial. J Am Coll Cardiol 22:955-962, 1993
7. Farr NL: Questioning placebo controls Science 288:1747, 2000 (letter)

In Reply: We, too, are concerned with the possibility of selection bias and lack of standardized data collection when making historical comparisons. That is why we stated, "The historical data required could be the survival or progression-free survival experience for a group of patients with the same stage of disease and amount of prior treatment, similar organ function and performance status, and for whom the same procedures were used for monitoring disease progression. Preferably, this historical experience would come from patients treated at the same institutions with the same referral patterns in a recent era, so that similar diagnostic methodologies and supportive care were available."[1] In clinical situations in which such data do not presently exist, we noted one possibility is to acquire the data prospectively in ongoing trials.

When the required historical data are not presently available and an agent is ready to be tested, we recommended performing a small screening randomized trial or a large definitive randomized trial and gave some criteria for choosing between the two. Drs Ratain and Stadler are concerned that using screening trials will result in a false-positive rate of 20%, and these false positives will need to be followed up with larger definitive trials. By definition, a screening trial will have more false positives than the 5% we would expect with a definitive trial. However, if the choice is between testing 20% of the agents in large trials or 100% of the agents in large trials, we believe that the oncologic clinical research environment is less likely to support testing 100% of the agents. With many agents currently under development, hard choices of which agents to test in definitive trials will have to be made. We believe the data from randomized screening trials and one-armed trials using valid histor-

ical comparisons will help to prioritize which agents should be tested first, especially if these trials can incorporate correlative studies that may confirm effects of the agents on their putative targets.

Ratain and Stadler are particularly concerned by our general recommendation against the randomized discontination (enrichment) trial design. Let us consider in more detail the ongoing Cancer and Leukemia Group B trial in metastatic renal cancer mentioned by them. This trial involves the accrual of as many as 335 patients and will test whether an 8-month treatment of carboxyamidotriazole (CAI) leads to longer maintenance of stable disease than a 4-month treatment of CAI among patients who have stable disease after 4 months on the treatment.[2] First, note that the required relatively large sample size limits the utility of this type of trial design for screening a large number of agents. Second, if this trial has a positive result, as we all hope, then we will know CAI has some efficacy in treating renal cancer. However, its efficacy in patients who have not already been treated with the agent for 4 months with stable disease will not be known. This may make the indication for use of the agent problematic, especially if only a small proportion of patients treated with CAI have stable disease at 4 months. In addition, it is not obvious what follow-up trials could be performed to clarify this situation after a positive result. However, our major concern with this trial is if it shows there is not a large or statistically significant difference between 8 versus 4 months of treatment for those patients with stable disease. Ratain and Stadler suggest that, in this case, the agent is highly unlikely to have a significant effect on disease progression or patient survival. We know of no evidence on this point and are hesitant to eliminate the development of agents based solely on a negative trial of this sort. Even so, as we noted previously, we could recommend an enrichment design when it is believed to be impossible to conduct a trial with a standard design. In this particular instance, the investigators stated that this was the case, and this was in part the reason that the Cancer Therapy Evaluation Program is sponsoring the trial.

Finally, we also encourage the development of new trial designs and nontraditional sequences of types of trials when needed. In fact, the point of our article was to discuss some of the options and their limitations for cytostatic agents. However, Ratain and Stadler seem also to suggest that the paucity of success to date in developing these agents is because of the use of standard trial designs. We would suggest that most of the reason for any lack of success to date has been a result of the ineffectiveness of the agents tested.

Edward L. Korn
Susan G. Arbuck
James M. Pluda
Richard Simon
Richard S. Kaplan
Michaele C. Christian
National Cancer Institute
Bethesda, MD

REFERENCES

1. Korn EL, Arbuck SG, Pluda JM, et al: Clinical trial designs for cytostatic agents: Are new approaches needed? J Clin Oncol 19:265-272, 2001

2. National Cancer Institute: Cancer Trials. Http://cancertrials.nci.nih.gov

Does Detection of Circulating ONYX-015 Genome by Polymerase Chain Reaction Indicate Vector Replication?

To the Editor: We wish to state that we find the conclusions of the recent publication by Nemunaitis et al[1] to be overstated. In particular, we take exception to the statement, "Detectable circulating ONYX-015 genome suggestive of intratumoral replication was identified in. . . ." We also disagree with the statement that ". . . persistent detection in two patients 10 days after the last injection suggests that a viral replication process was ongoing,. . . ." In our analysis of over 190 subjects treated intratumorally with RPR/INGN 201, a replication-defective adenovirus carrying the human *p53* gene, we have seen clearance of the viral genome similar to that reported by the authors, with detectable vector-related sequences found in urine up to 28 days after the last injection. This was observed over multiple treatment cycles. Preliminary data from one of our phase II studies, showing the presence of vector DNA in plasma for 13 days after a single intratumoral administration at a dose of 1.0×10^{12} virus particles and for at least 15 days after the last of six administrations (1.0×10^{12} virus particles per dose), were presented at the Annual Meeting of the American Society of Clinical Oncology in 1999.[2] The data from cycle 1 of the completed study are summarized in Fig 1 for each of the two schedules of administration (one administration [schedule A] or six administrations spaced over 2 weeks [schedule B]).

Using Taqman (Applied Biosystems, Foster City, CA) real-time polymerase chain reaction evaluation of samples from multiple sources, as the above authors did, we were able to demonstrate peak values of the administered product at 1 to 2 days after injection, with levels dropping to undetectable after 17 to 20 days after the first injection (Table 1).

However, measurable levels of RPR/INGN 201 sequences were routinely detectable at day 10 after injection in numerous subjects. Thus, we regularly detect shedding of viral DNA (though not necessarily of virus), much like that reported by the authors.

These samples were also subject to analysis for the presence of virus able to cause cytopathic effect (CPE) on the 293 and A549 cell lines. Although CPE was often observed in samples tested on 293 cells (permissive for replication of E1A-deleted adenovirus constructs such as RPR/INGN 201), only two cases subsequently identified as adenovirus type 11 gave CPE on A549 cells. Thus, we see low levels of shedding that can only be attributed to a replication-defective virus. In addition, preclinical studies using RPR/INGN 201 have been unable to detect virus replication other than in 293 cells.

In 27 cases, viral isolates from study subjects were amplified by several rounds of blind passage on 293 cells, allowing for replication of RPR/INGN 201. Of these, 26 successfully amplified and were subject to Southern Blot analysis. In each case, no rearrangements of the RPR/INGN 201 genome were detected, indicating a high degree of genetic stability. Thus, where virus could be recovered from several subjects and amplified by serial passage, it remained replication-defective.

We believe that it is important to point out that persistence of shedding viral sequences, or of virus, does not a priori demonstrate that viral replication is taking place. Indeed, our observations of similar biodistribution of a replication-defective adenovirus (RPR/INGN 201), in the absence of any detectable replication-competent adenovirus, lead us to the opposite conclusion, namely that the use of the Taqman real-time polymerase chain reaction technology is

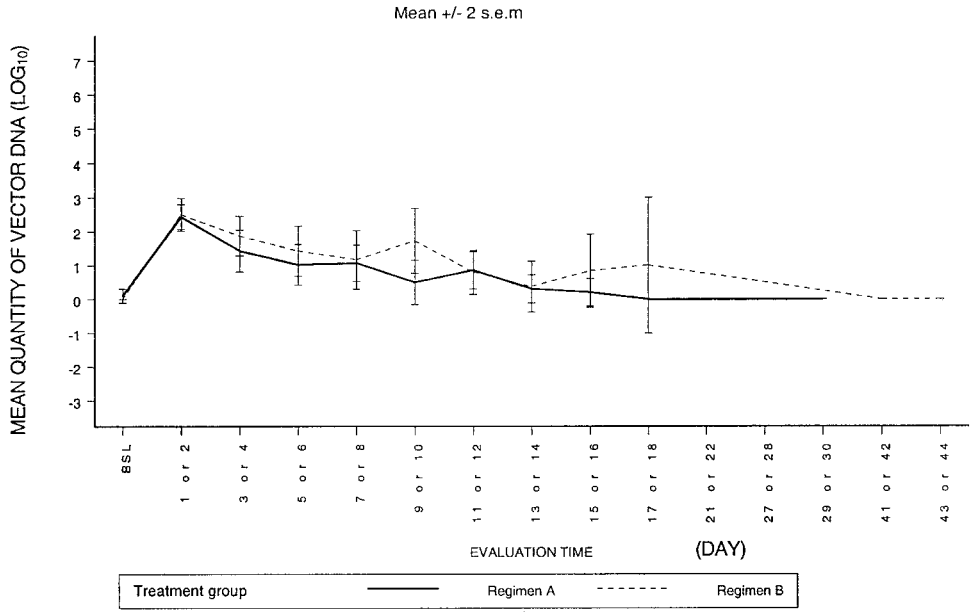

Fig 1. Plasma pharmacokinetics of RPR/INGN 201: plot of mean quantity of vector DNA in plasma as a function of time by treatment group based on quantitative polymerase chain reaction analysis (cycle 1 only, all treated patients).

Table 1. Samples Positive for RPR/INGN 201 Polymerase Chain Reaction at Baseline or During Cycles, by Sample Type

	No. of Samples Positive/Total Samples				
	Cycle 1		Cycle 2		≥ Cycle 3
	Baseline	Any Day	Pretreatment	Any Day	Any Day
Feces	0/6	3/34	0/1	4/21	0/8
Gargle	0/35	86/195	1/12	32/78	7/14
Lymphocytes*	4/17	49/101	5/5	22/50	6/11
Plasma	1/37	105/203	1/12	42/84	2/15
Urine	1/34	13/199	1/11	10/79	0/14

*Peripheral blood lymphocytes.

not sufficient to prove that viral reproduction is taking place, as suggested by the authors.

We appreciate the opportunity to present an alternative point of view. It is expected that this work will be presented in detail at this year's meeting of the American Society of Clinical Oncology and will then be published in extenso. Portions of it have also been presented at the 1999 American Association of Cancer Research meeting[3] and by Clayman et al.[4]

Antoine Yver
Aventis Pharmaceuticals Inc
Bridgewater, NJ

REFERENCES

1. Nemunaitis J, Khuri F, Ganly I, et al: Phase II trial of intratumoral administration of ONYX-015, a replication-selective adenovirus, in patients with refractory head and neck cancer. J Clin Oncol 19:289-298, 2001

2. Constenla-Figueiras M, Betticher C, Del Campo J, et al: A phase II trial with Ad5CMV-*p53* as a single agent in recurrent/refractory SCCHN looking at vector biodistribution and horizontal transmission under normal life conditions. Proc Am Soc Clin Oncol 18:444a, 1999 (abstr 1714)

3. Constenla-Figueiras M, Fey M, Del Campo J, et al: A phase II trial with Ad5CMV-*p53* as a single agent in recurrent/refractory SCCHN looking at vector biodistribution and horizontal transmission under normal life conditions. Proc Am Assoc Cancer Res 40:189, 1999 (abstr 1261)

4. Clayman G, el-Naggar A, Lippman S, et al: Adenovirus-mediated *p53* gene transfer in patients with advanced recurrent head and neck squamous cell carcinoma. J Clin Oncol 16:2221-2232, 1998

In Reply: We appreciate the additional unpublished data provided by Dr Yver, particularly as published literature has not identified detection ($> 1 \times 10^3$ plaque forming units/0.5 mL) of RPR/ING201 in blood for more than 90 minutes.[1-3] Nevertheless, although we agree with Yver that injection of replication-defective adenoviral vectors may be associated with transient circulation of viral genome, we still maintain that the prolonged detectable circulation of Onyx-015 genome as detailed in our article is suggestive of intratumoral replication.[4] We agree that our data does not prove replication. However, much of our argument is based on the consideration that, given the assay's sensitivity, judgements regarding the significance of polymerase chain reaction detection of nucleic acids always hinge strongly on the precise definition of what constitutes a threshold detection limit. Our assay was validated for a detection limit sensitivity to 1.05×10^4 Ad particles/mL, with a qualification sensitivity of 4.2×10^4 Ad particles/mL. Our reported results are all above the sensitivity level (ie, $> 4.2 \times 10^4$ particles/mL), and the assay was validated to be specific for

Onyx-015 and was performed independently by Althea Technologies (San Diego, CA). Because Yver has not published his results, it is difficult to interpret his statement. Once his data are published, the scientific community can rigorously evaluate the sensitivity and specificity of his assays and data in the context of ours. Further, although Yver's methods of demonstrating that the viral genome recovered was intact virus are elegant, we also note that E1-deleted vectors such as RPR/ING201 can, and do, replicate in some cancer cells.[5,6] This is a critical issue to determine whether or not Yver's observation represents viral replication in his system and one that cannot be evaluated on unpublished data.

In addition, in one of our subsequent trials where we administered intravenous Onyx-015, we demonstrated plasma levels of Onyx-015 genome between 5.6 and 890 $\times$ 10^4 genomes/mL 7 days after a single infusion (2×10^{10} to 2×10^{13} particles/mL). And further, we were able to demonstrate an increase in circulating viral genome at 48 hours 2.5- to 10-fold above those seen 6 hours after initial Onyx-015 infusion. Such an increase with already significant levels of circulating viral genome is most easily explained by continued viral replication. Finally, other evidence provided by electron micrographs of biopsied malignant tissue reveal bulging nuclei filled with Onyx-015 during the period when viral genome is detected in circulation, whereas adjacent nonmalignant tissue shows no evidence of viral presence.[7,8]

In summary, we reported the first peer-reviewed data of persistently detectable viral genome after administration of replication-competent virus, which does suggest viral replication. This result is based on rigorously validated assays in the context of peer-reviewed publication. However, we appreciate Yver's point that further research is needed. We are in the process of development of a quantitative plaque assay to potentially measure functional Onyx-015 from stored plasma samples of previously treated patients with detectable circulating genome. We thank Dr Yver and the editors for the opportunity to expand here the discussion of this important area.

John J. Nemunaitis
Casey Cunningham
US Oncology
Dallas, TX

REFERENCES

1. Clayman G: Adenovirus-mediated p53 gene transfer in patients with advanced recurrent head and neck squamous cell carcinoma. J Clin Oncol 16:2221-2232, 1998

2. Nemunaitis J, Swisher G, Timmons T, et al: Adenovirus-mediated p53 gene transfer in sequence with cisplatin to tumors of patients with non-small-cell lung cancer. J Clin Oncol 18:609-622, 2000

3. Swisher SG, Roth JA, Nemunaitis J, et al: Adenovirus-mediated p53 gene transfer in advanced non-small-cell lung cancer. J Natl Cancer Inst 91:763-771, 1999

4. Nemunaitis J, Khuri F, Ganly I, et al: Phase II trial of intratumoral injection of ONYX-015, a replication-selective adenovirus, in patients with refractory head and neck cancer. J Clin Oncol 19:289-298, 2001

5. Steinwaerder D, Carlson C, Otto D, et al: Tumor-specific gene expression in hepatic metastases by a replication-activated adenovirus vector. Nat Med 7:240-244, 2001

6. Steinwaerder D, Carlson C, Lieber A: DNA replication of first-generation adenovirus vectors in tumor cells. Gene Therapy (in press)

7. Nemunaitis J, Cunningham C, Buchanan A, et al: Intravenous infusion of a replication-selective adenovirus (ONYX-015) in cancer patients: Safety, feasibility and biological activity. Gene Therapy (in press)

8. Nemunaitis J, Ganly I, Khuri F, et al: Selective replication and oncolysis in p53 mutant tumors with ONYX-015, an E1B-55kD gene-deleted adenovirus, in patients with advanced head and neck cancer: A Phase II trial. Cancer Res 60:6359-6366, 2000

Assessing Adjuvant Breast Cancer Therapy Benefit

To the Editor: The article by Loprinzi and Thomé[1] on adjuvant systemic therapy for primary breast cancer offers minimal assistance in deciding whether to give adjuvant chemotherapy in breast cancer. There is a general consensus that patients with node-positive disease (especially estrogen receptor–negative) and large tumors benefit from chemotherapy. The major problem arises in dealing with the increasingly frequent mammogram-discovered, small, node-negative lesion. The authors' approach is to take the collected data from the Early Breast Cancer Trialists articles[2,3] (which stratify patients by presence or absence of nodes but not by tumor size) and assume that the annual proportional reduction in risk applies equally to all tumors of a given nodal and estrogen receptor status, irrespective of size. This information is then fed into an opinion-generated table of baseline risk as a function of tumor size and nodal status to generate tables of clinical benefit (Tables 7 and 8).[1]

There are a number of problems with this approach: (1) The input baseline risk (Table 2) is only opinion generated. Correlation is claimed with Dr Peter Radvin's "*Aduvant!*" program based on Surveillance, Epidemiology, and End-Results data and a proprietary formula (Fig 1). However, close inspection of that figure in the low-risk, high-survival range shows significant differences of magnitudes similar to the claimed benefits of chemotherapy. (2) The assumption that tumors of different sizes have the same annual proportional reduction in risk is questionable. The data analyzed by the Early Breast Cancer Trialists are not stratified by tumor size. Furthermore, the data are based on studies a decade old that include mostly patients with larger tumors or node-positive tumors. Extrapolation to small node-negative tumors could well be in error. Perhaps there are threshhold effects below which metastasis is exceedingly unlikely. (3) Table 8 is filled with rows of entries for several different ranges of involved nodes. These data are not particularly useful because all of these patients are treated anyway. The issue is the small node-negative tumors where the data are most suspect. (4) The patients in question with small node-negative tumors have low recurrence rates. Other variables not included in this model, such as histology, could have large relative effects. (5) The information on paclitaxel, based upon a single abstract, has now become highly suspect.[4]

As a medical oncologist I am still left with the problem of whether to recommend treatment of women with small, node-negative tumors at considerable toxicities and with minimal statistical benefit. Certainly the marginal statistical benefit implies that we are subjecting many more women to toxicity than will benefit from therapy. The answer will come from better prognostic markers, such as Braun et al's[5] work with bone marrow markers or perhaps ultimately from gene expression

arrays.[6] More effort needs to be directed in these areas rather than continuing to rehash old data.

Arthur M. Sleeper
Ravenel Oncology Center
Martinsville, VA

REFERENCES

1. Loprinzi CL, Thomé SD: Understanding the utility of adjuvant systemic therapy for primary breast cancer. J Clin Oncol 19:972-979, 2001

2. Early Breast Cancer Trialists' Collaborative Group: Polychemotherapy for early breast cancer: An overview of the randomised trials. Lancet 352:930-942, 1998

3. Early Breast Cancer Trialists' Collaborative Group: Tamoxifen for early breast cancer: An overview of the randomised trials. Lancet 351:1451-1467, 1998

4. National Institutes of Health: Adjuvantive therapy for breast cancer. NIH Consensus Statement Online 2000, November 1-3, 17:1-23, 2000

5. Braun S, Cevatli BS, Assemi C, et al: Comparative analysis of micrometastasis to the bone marrow and lymph nodes of node-negative breast cancer patients receiving no adjuvant therapy. J Clin Oncol 19:1468-1475, 2001

6. Hedenfalk I, Duggan D, Chen Y, et al: Gene-expression profiles in hereditary breast cencer. N Engl J Med 348:539-548, 2001

To the Editor: In the article by Loprinzi and Thomé,[1] the authors describe, as an example, a 55-year-old, node-negative, estrogen receptor–positive breast cancer patient who will improve her 10-year disease-free survival from 70% to 79% with combined adjuvant therapy while her chance of being free of disease recurrence will actually improve to 87%. In other words, she apparently benefits an additional 8% because she will die of another cause. Although at first glance these distinctions are confusing, such calculations underscore the obvious importance of considering age and comorbidity in advising patients and also in realizing that although the *Numeracy* program, according to the authors, "can factor in the mortality that would be associated with a woman who was otherwise healthy, based on causes of mortality other than breast cancer," the online program currently available (http://mhs.mayo.edu/adjuvant) does not.

Steven M. Sorscher
Leah L. Dietrich
Gundersen Clinic, Ltd
La Crosse, WI

REFERENCE

1. Loprinzi CL, Thomé SD: Understanding the utility of adjuvant systemic therapy of primary breast cancer. J Clin Oncol 19:972-979, 2001

In Reply: We appreciate the opportunity to respond to the letters authored by Dr Arthur M. Sleeper and Drs Steven M. Sorscher and Leah L. Dietrich. First, Sorscher and Dietrich insightfully raise a seemingly confounding issue dealing with why the overview analysis reports an improved outcome (in terms of annual proportional reductions in risk recurrence) for disease-free survivals versus overall survivals.[1] This reported differential effect suggests, in the patient example cited, that combined adjuvant therapy provides an 8% additional benefit for survival (as opposed to disease-free survival). There are at least three reasons for this discrepancy, one of which, as Sorscher and Dietrich surmise, is that some patients, despite developing recurrent breast cancer, will die of causes other than breast cancer. Two other reasons, however, are that (1) the overview analysis counts a contralateral breast cancer as a recurrence and (2) the delay between recurrence and death makes the benefits sound more robust regarding recurrence, as opposed to death. As might be predicted, the latter reason should become less prominent as the overview data further mature. This effect is evident in the most recent overview analyses, which were recently presented at a recent National Institutes of Health Consensus Development Conference but which have yet to be published.

It is noteworthy, as we discussed in our publication, that we use the terms 10-year survival and 10-year disease-free survival interchangeably when talking about eventual prognosis, given that these two percentages are relatively close to each other. It is true that the online version of *Numeracy*, at http://mhs.mayo.edu/adjuvant, does not allow for factoring in comorbid conditions. For this online version of *Numeracy*, we felt that a simple pull-down menu with automated data entry would be best for most expected users. Nonetheless, there is an available *Numeracy* spreadsheet program that does have the ability to change 10-year baseline risk estimates and efficacy estimates at will. We are happy to provide interested physicians with a copy of this spreadsheet program (please request at email: cloprinzi@mayo.edu).

Moving on to respond to comments from Dr Sleeper, let us address the five problems that he lists. Sleeper first takes issue with the baseline risk that we used. We admit that this risk is opinion generated. However, it is not individual physician opinion-generated information, which, as illustrated in Table 2 of our article[2] and in a previous publication,[3] is very diverse. Rather, it is an average opinion from a number of clinicians who regularly see patients with breast cancer. We are impressed with the remarkable correlation between prognosis seen in our program with that seen with the method from Ravdin et al.[4] With regard to the relatively modest differences in opinions between our baseline estimates and those of Ravdin et al,[4] it is not clear which one is more or less correct. Although we are hopeful that more precise baseline prognoses will be generated in the future, we are unaware of more helpful baseline prognosis information to use for individual patients.

Sleeper's next concern deals with whether annual proportional risk reduction is an accurate way of understanding chemotherapy effects across patients with different tumor sizes and different nodal situations. We note that we did not define this relative or proportional mechanism of describing the benefits of adjuvant therapy, a mechanism that has been routinely used to describe adjuvant treatment benefits in recent reports of clinical trials and in adjuvant meta-analyses and overviews. Rather, our article sought to better translate individual patient proportional benefits into absolute benefits.

Sleeper's next concern deals with Table 8. In contrast to Sleeper's assertion, this table contains information regarding both patients with and without involved axillary lymph nodes. We agree with Sleeper that the information regarding small, node-negative cancers might be most helpful for allowing a patient and her physician to better understand the potential benefits from systemic therapy. In addition, however, we believe that understanding the relative benefits of adjuvant therapy in node-positive patients would be helpful for many patients so that they can better know the magnitude of benefit for which they are receiving toxic therapy.

The fourth problem that Sleeper notes is that other prognostic variables, such as histology, were not included in our baseline prognostic model. We will admit that there are many different prognostic

tools that have been studied in groups of patients. Nonetheless, given the vast diversity of opinions regarding the best two established features (that being the number of involved axillary lymph nodes and tumor size), the use of less robust prognostic factors, in individual patient cases, is not established. Sleeper is referred to Fig 4 in a previously published article,[2] which illustrates the limited value from adding weaker prognostic factor information to prognostic opinions based on axillary lymph node status and tumor size.

Sleeper's last concern deals with the fact that paclitaxel is not fully established, effective therapy for patients in the adjuvant setting. We agree, as indicated by reminding him that this was noted in our article on four separate occasions.

Sleeper ends his letter looking to the future when, hopefully, better prognostic information will be available, as opposed to trying to better understand how to use the data that are accessible to us today. We, too, look forward to the future when better prognostic information should be available. In contrast to Sleeper, however, we do believe that there is room for a better understanding of the prognostic information that is currently available so as to better be able to inform and treat those patients we see in our practice today. The multiple, positive, unsolicited comments that we have received from colleagues since the publication of this article indicates to us that compiling available information in an accessible manner is worthwhile. Sometimes old hash does taste better after rewarming.

Charles L. Loprinzi
Stephan D. Thomé
Mayo Clinic
Rochester, MN

REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group: Polychemotherapy for early breast cancer: An overview of the randomised trials. Lancet 352:930-942, 1998
2. Loprinzi CL, Thomé SD: Understanding the utility of adjuvant systemic therapy for primary breast cancer. J Clin Oncol 19:972-979, 2001
3. Loprinzi CL, Ravdin PM, De Laurentiis M, et al: Do American oncologists know how to use prognostic variables for patients with newly diagnosed primary breast cancer? J Clin Oncol 12:1422-1426, 1994
4. Ravdin PM, Siminoff LA, Davis GJ, et al: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 19:980-991, 2001

Measurement of Brain Tumor Volumes by the Perimeter Method

To the Editor: The careful analysis by Sorensen et al[1] on measurement of brain tumor volumes by the perimeter method clearly establishes this to be a more accurate and probably simpler method of brain tumor volume measurement. The authors consider a partial response to be a 50% or more decrease in tumor volume.

However, for decades a 50% decrease in tumor cross-sectional area has been the standard used by medical oncologists, including the World Health Organization Handbook[2] that Sorensen cited, for determining that a tumor has reached a partial response. True, measurement of glioblastoma shrinkage by careful perimeter volume measurements may be reasonable and more accurate for that kind of tumor. However, this has not been and should not be standard for measurement for tumors elsewhere in the body, because the vast oncology literature on tumor response to chemotherapy has been based on a 50% reduction in area, and area measurements are easily performed by the oncologist on chest x-rays, computed tomography scans, and lymph nodes.

The volume criterion for response is less stringent than area; a 50% reduction in tumor volume represents only a 37% reduction in cross-sectional area. Conversely, to reach the 50% reduction in area, a tumor has to reduce in volume by approximately 65%. If the oncology community is going to change from using cross-sectional area to volume as the standard for tumor shrinkage in describing responses to chemotherapy, it should be done explicitly by international consensus and not by nibbling away at it, disease by disease.

Charles M. Bagley, Jr
Puget Sound Cancer Centers
Seattle, WA

REFERENCES

1. Sorensen AG, Patel S, Harmath C, et al: Comparison of diameter and perimeter methods for tumor volume calculation. J Clin Oncol 19:551-557, 2001
2. World Health Organization: Handbook for Reporting Results of Cancer Treatment. Geneva, Switzerland, World Health Organization, 1979

In Reply: We appreciate Dr Bagley's interest in our article. We agree with Bagley that a 50% change in volume differs from a 50% change in cross-sectional area, and we concur that a 50% decrease in cross-sectional area has been a standard for response definition in the past. In our article, the choice of a 50% change in volume (rather than, for example, a 65% reduction in volume) was somewhat arbitrary and chosen for convenience only. Indeed, our results demonstrating the potential impact of the greater variability of the cross-sectional method compared with the volumetric method remain intact and relevant no matter what threshold is chosen to determine response.

We purposely hoped to avoid the ongoing discussion (some might suggest controversy) surrounding how the oncology community chooses to define various response parameters. Our aim was instead to illustrate that efficient methods now exist to reduce the variance of these measures, and that this reduction in variance is large enough to improve estimates of tumor response and thereby potentially allow smaller patient groups to be used to test novel therapies. This reduction in variance is not a result of choosing a volumetric method instead of a cross-sectional method, but rather a result of computer assistance in the segmentation.

However, Bagley may imply in his final comments about the accessibility of film images to the oncologist that the simpler measurement is also preferable in clinical trials. We believe one of the implications of our results is that an oncology trial that uses semi-automated tumor quantitation could require significantly fewer subjects, and that this added benefit could offset the issue that fewer investigators could perform the analysis on-site. We thank Dr Bagley for agreeing that this reduction in variance seems useful and

apologize for any confusion that our choice of measurement change may have caused.

A. Gregory Sorensen
Massachusetts General Hospital
Charlestown, MA

Prostate-Specific Antigen Levels in Nipple Aspirate Fluid

To the Editor: We reviewed with interest the article by Zhao et al.[1] The authors collected nipple aspirate fluid (NAF) from women with recently diagnosed cancer, those with atypia or in situ carcinoma, and those without breast lesions. They were successful in collecting NAF from 34% of subjects and had adequate NAF to perform analyses in 29.6% of subjects who consented to undergo aspiration. The authors observed similar levels of prostate-specific antigen (PSA) in tumor-free breasts and in those recently diagnosed with breast cancer. The authors speculate that the differences observed in our original study[2] were a result of hemodilution in the NAF collected from mastectomy specimens, which, in turn, resulted in lower PSA levels.

We have obtained NAF from 97% of the last 500 women who consented to undergo the procedure. We divided the NAF specimen equally, half for cytologic review and half for other biomarker studies. One of the aspects of cytologic review is to evaluate the specimen for red cells. We recently evaluated specimens from 110 women with breast cancer. Thirty-two specimens (29%) were observed to contain red blood cells. There was no statistically significant difference in NAF PSA levels in specimens with and without blood whether considering the population as a whole or grouped by menopausal status. When we eliminated specimens containing blood and compared PSA levels in women with and without breast cancer, controlling for menopausal status and age, we found that PSA was significantly lower in NAF from the breasts with cancer than from the breasts without cancer ($P = .0001$).

We are not sure why the results of the two studies differ, but we do not believe it is because of hemodilution. One difference was the fraction of women from whom a sample was obtained. The authors do not indicate the fraction of subjects from whom NAF was collected in the tumor-free versus tumor-containing breasts. Because NAF was obtained from only one third of the subjects consenting to undergo the procedure, the population studied may not reflect the population as a whole.

Many of our NAF specimens from women with breast cancer have come from mastectomy specimens. We perform aspiration immediately after the breast is removed from the chest wall. In the article by Zhao et al,[1] NAF was performed in women after they had been diagnosed with breast cancer but before definitive therapy had been instituted. We elected to forego aspiration on women with proven breast cancer before mastectomy both because the breast is often painful after diagnostic needle biopsy and because of the subject's heightened anxiety. Fortunately, the weaknesses of both approaches are being addressed by us and by other groups prospectively performing nipple aspiration in women scheduled to undergo surgery for a suspicious lesion without a definitive diagnosis. These studies have a number of benefits. They minimize potential bias due to a diagnosis of breast cancer. Also, because the subject has not undergone a fine- or core-needle biopsy that leaves her breast tender, often too tender to undergo nipple aspiration, one is more likely to obtain NAF. These studies should help determine the usefulness of PSA levels in breast aspirate fluid in predicting the chance that a subject has or will develop breast cancer.

Edward R. Sauter
Thomas Jefferson University
Philadelphia, PA

Eleftherios P. Diamandis
Mt Sinai Hospital
University of Toronto
Toronto, Ontario, Canada

REFERENCES

1. Zhao Y, Verselis SJ, Klar N, et al: Nipple fluid carcinoembryonic antigen and prostate-specific antigen in cancer-bearing and tumor-free breasts. J Clin Oncol 19:1462-1467, 2001

2. Sauter ER, Daly M, Linahan K, et al: Cancer Epidemiol Biomarkers Prev 5:967-970, 1996