Cancer Chemother. Pharmacol. 3, 1–5 (1979)

Cancer Chemotherapy and Pharmacology

Editorial

Methodology of Data Reporting in Advanced Breast Cancer Trials

S. K. Carter

Northern California Cancer Program,
1801 Page Mill Road, Bldg B, Suite 200, Palo Alto, California 94304, USA

A major problem in the oncologic literature is the lack of comparability in definitions and data reporting techniques. In attempting to answer important questions from the literature, there appears to be a great deal of data available. The superficial impression is that answering the question should not be difficult. If reading is restricted to the abstract, and perhaps selected tables and the discussion section, easy answers will still seem possible. Clinical research papers tend to have relatively short Methods sections in comparison to research papers from the laboratory. If the Methods sections of clinical research papers are examined closely, then facility of interpretation becomes more problematical. One of the major problems is that there is no comparability as to what is included in the Methods section. In addition, the criteria of response used are poorly defined, with minimal detail in critical areas of subtlety and difficulty. A case in point is the treatment of advanced breast cancer with chemotherapy. In 1978, the journal *Cancer* published six papers on the treatment of advanced breast cancer with combination chemotherapy regimens involving adriamycin (Table 1). These six papers came from one major American cooperative group, i.e., Southwest Oncology Group (SWOG), four major American cancer centers (Bowman Gray University, National Cancer Institute, Roswell Park Memorial Institute, and Yale University), and one English cancer center (Guys Hospital). It is the purpose of this paper to review these six papers as to their definitions of response and other important criteria included in their Methods section. Since all were published in the same major journal and came from major institutions they are presumably representative of the current state of comparability, or lack thereof, in our current literature.

One of the first aspects of attempts to compare results is the consideration of critical areas of case selection. In Table 2, the six papers are compared for five major selection variables in their study. What stands out is the lack of any information on certain critical variables. Three papers do not mention any minimum white blood cell count and platelet count for inclusion. Does this mean that any blood count could be included, with appropriate dose reduction? If so, this would give a different mix of patients than in the two cooperative group studies where near-normal counts were required for entry. The Guys Hospital trial, on the other hand, included patients with some initial depression of blood counts. Since adriamycin was involved in all the six studies, it would be important to know what exclusions were made for prior cardiac disease. As can be seen, this was not mentioned in half the papers. One of the most important variables that has been shown to affect chemotherapy response and survival is the performance status. It is worth noting that in only one of the papers was an exclusion for performance status mentioned. Is one to assume therefore that all performance statuses were theoretically acceptable for entry into the study? In the papers from Bowman Gray University, Roswell Park, and Yale University, it is obvious that this was so if one looks at the tables describing the pretreatment characteristics where all performance statuses were broken down.

In the chemotherapy of breast cancer today, most of the objective responses observed are less than complete. Because of this, it is of critical importance to define a 'partial response'. Defining a partial response is an area of great complexity, despite the fact that > 50% shrinkage of measurable lesions is now an accepted generality. Table 3 compares the definition of partial response in the six papers. The definitions in Table 3 are taken verbatim from the Methods section, and are complete. Three areas are covered. The first is the general definition of a partial response. The second area includes statements about bone lesions, and the third, statements about liver lesions.

Regarding the general definition of a partial response, all the six papers look superficially comparable since all utilize the > 50% shrinkage with measurement

0344-5704/79/0003/0001/$ 01.00

Table 1. Six breast cancer papers published in the journal Cancer in 1978

Group	Regimens		
	I	II	III
Southwest Oncology, Group [8]	Adriamycin + cyclophosphamide	Adriamycin + cyclophosphamide + 5-FU	Adriamycin + cyclophosphamide + 5-FU + methotrexate
Bowman Gray University [6]	Cyclophosphamide + methotrexate + 5-FU + vincristine + prednisone	Cyclophosphamide + adriamycin + 5-FU + vincristine + prednisone	
National Cancer Institute [2]	Cyclophosphamide + methotrexate + 5-FU	Cyclophosphamide + adriamycin + 5-FU	
Roswell Park [5]	Cyclophosphamide + 5-FU + prednisone	Adriamycin	Adrenalectomy
Yale University [4]	Adriamycin + cyclophosphamide	Adriamycin + cyclophosphamide + 5-FU + methotrexate	
Guys Hospital [7]	Adriamycin + vincristine	Adriamycin + vincristine + norethisterone acetate	

Table 2. Comparison of five potential exclusion criteria in six breast cancer papers

Group	Prior chemotherapy	Minimum WBC allowed	Minimum platelets allowed	Cardiac exclusions	PS
SWOG	Excluded	4,000	100,000	1. Known heart disease 2. Abnormal EKG, except non-specific ST-T wave changes	NM
Bowman Gray University	Excluded if exposed to drugs in protocol	4,000	120,000	NM	NM
NCI	Excluded	NM	NM	NM	KS > 30
Roswell Park and Albany	Excluded	NM	NM	NM	NM
Yale University	Single-agent chemotherapy other than adriamycin included	NM	NM	Moderate to severe cardiac disease that might preclude the administration of adriamycin excluded	NM
Guys Hospital	Probably excluded, although not specifically mentioned in methods section	2,000	> 70,000	Cardiac function 'normal'	NM

of two perpendicular diameters of lesions. What is difficult to elucidate is how a partial response is defined with multiple lesions. As can be seen in Table 4, there are four general possibilities for defining a partial response with multiple lesions, utilizing the 50% shrinkage criterion. These are distinctly not comparable and can lead to differing partial response rates with the same clinical material. An example is shown in Table 5, where the measurements of three lesions are shown. These lesions measure 10 × 10, 3 × 3, and 4 × 4. If 8 weeks later the

Table 3. Partial response criteria in six recent breast cancer papers

Group	General definition of PR	Bone lesions	Liver lesions
SWOG	A reduction of at least 50% in the product of the two widest diameters of measurable lesions	Reduction of bone pain, and as either a partial or complete response depending on the degree of healing as evidenced by X-ray	NM
Bowman Gray University	At least a 50% reduction of the product of the longest perpendicular diameter of measurable lesions. No corresponding increase in other areas of known disease	Improvement of bone lesions on scan or skeletal X-ray for at least 3 months, in addition to significant subjective improvement	A 30% or greater decrease in the sum of the measurements below both costal margins at the midclavicular line and xiphoid process, with stabilization or improvement in LFTS
NCI	At least a 50% reduction in the sum of the products of the longest perpendicular cross-sectional measurements in the involved organ systems – reduction had to be present in 50% or more of the involved organ sites for at least 2 measurement periods separated by at least 4 weeks	A partial recalcification and mixed osteoblastic and osteolytic lesions not considered evaluable	A 30% reduction – sum of the measurements in the right and left midclavicular and xiphoid lines
Roswell Park and Albany	At least 50% reduction in the product of the two perpendicular diameters of *all* measurable lesions in the absence of progression of other lesions	NM	NM
Yale University	A reduction of at least 50% of the product of the largest perpendicular diameters of measurable lesions, without the appearance of new areas of measurable disease for at least 1 month	A decrease in any bone pain present required. Regression of osteoblastic lesions or recalcification of osteolytic lesions	NM
Guys Hospital	A > 50% decrease in the sum of the products of the perpendicular axis of measurable lesions and objective improvement in evaluable but non-measurable lesions; no new lesions. It was not necessary for every lesions to have regressed	NM	NM

10 × 10, lesion has shrunk to 2 × 2 and the other two lesions have remained stable, is this a partial response? If a 50% decrease in the sum of the products of all measurements of lesions is used as the criterion, this is a partial response as the sum of the products has declined from 125 to 29. On the other hand, if shrinkage of all lesions, or even half the lesions, is the criterion then this is not a partial response. If, 8 weeks later, the two smaller lesions have disappeared but the large lesion has remained unchanged then the situation reverses. On the basis of the sum of products, this is not a partial response, but on that of shrinkage of more than half the lesions, it is. In both cases, the strict criterion of all lesions shrinking would exclude a partial response definition. In both cases the liberal definition of only one lesion having to shrink would include this as a partial response. It should therefore be required that partial response criteria go into greater detail as to how a partial response is defined in the face of multiple lesions.

Both bone lesions and liver lesions are problems in advanced breast cancer, defying easy application in the > 50% shrinkage criterion. It is worth noting that in Table 2 two papers say nothing about how osseous lesions are evaluated and four say nothing about hepatic lesions. In addition, most of these papers say nothing about pleural effusions, CNS lesions, and lymphangitic pulmonary metastases, whether they are excluded, and if not, how they are evaluated.

In evaluating clinical trial results, the duration of response and survival are crucial aspects. Table 6 compares the methodology described for this in the six papers under review. As can be seen, only three papers describe the zero point for determination of response duration and only two do so for the determination of surviv-

Table 4. Possibilities for defining PR with > 50% shrinkage when multiple lesions exist

1. All lesions must shrink by 50% (product of perpendicular measurements)
2. More than half must shrink by 50%, while others remain stable
3. Only one must shrink by 50%, while others remain stable
4. 50% decrease in sum of products of all lesions

Other variables
Duration
Bone lesions
Liver enlargement
CNS lesions
Pleural effusions
Lymphangitic pulmonary metastasis

Table 5. Idealized situation with three lesions and selected shrinkage 8 weeks later

	Lesions measurement			Is this a 'PR'				
	1	2	3	Sum of the products	50% shrinkage age in sum of products	All lesions shrink 50%	Half the lesions shrink 50%	One lesions shrinks 50%
Base line	10 × 10	3 × 3	4 × 4	125	–	–	–	–
8 weeks later, example 1	2 × 2	3 × 3	4 × 4	29	Yes	No	No	Yes
8 weeks later, example 2	10 × 10	0 × 0	0 × 0	100	No	No	Yes	Yes

Table 6. Duration of response and survival criteria in six breast cancer papers

	Duration of response	Survival
SWOG	NM	NM
Bowman Gray University	Calculated from the first day of chemotherapy to time of progression or last date of contact	Comparisons made according to methods described by Wilcoxan and Gehan
NCI	NM	Calculated by life-table method of Kaplan and Meier and compared using the generalized Wilcoxan test of Gehan
Roswell Park and Albany	NM	Standard product-limit method of Kaplan and Meier and comparisons by two-tailed generalized Wilcoxan test
Yale University	Measured from the initiation of chemotherapy of the time of relapse	Calculated from initiation of chemotherapy until death or if patient still alive until the completion of the study on 3 January, 1977; actuarial method of Kaplan and Meier using the Gehan modification of the generalized Wilcoxan test
Guys Hospital	Dated from the start of chemotherapy until either new lesions appeared or any one existing lesion increased by 25% or more above its smallest recorded size	Dated from time of commencement of treatment until death, analyzed by life-table method

al. Only four papers mention the methodology used for developing the survival curves and making statistical comparisons for significance. Table 6 again displays verbatim the statements made in the Methods sections. Not being a statistician I can only wonder whether 'the generalized Wilcoxan test of Gehan' is the same as 'the Gehan modification of the generalized Wilcoxan test' or 'the two-tailed generalized Wilcoxan test'.

Discussion

In breast cancer, four types of parameters are potentially available for the measurement of response. The first type is bidimensionally measurable disease, i.e., pulmonary nodules on X-ray, subcutaneous or node masses. The second type is disease that is measurable but not bidimensionally. This would include liver en-

largement and intra-abdominal masses. A third type would be nonmeasurable disease such as lymphangitic pulmonary disease, ascites, and inflammatory skin lesions. While not measurable, these could be evaluable for response. The last type would be indirect parameters, i.e., hepatic enzymes or serum calcium evaluation. In patients with multiple skin or bone lesions it might be possible to have some lesions bidimensionally measurable while others were only unidimensionally measurable, and still others that were only evaluable and nonmeasurable. It is of critical importance in any study to specify exactly which type of lesions will allow for patients' entry and which will exclude. It is then essential to define clearly what will consitute a response for each type as regards shrinkage and its duration. Perhaps most critical of all is the need to determine what will be called a partial response in the face of multiple lesions of differing types.

Advances in cancer therapy can only come from the clinical investigation of new therapeutic approaches. If clinical research is successful then oncologic practice will be modified as a result. Critical to the incorporation of clinical research results into the practice of oncology is the ability to understand what the oncologic literature is saying. Superficial reading of abstracts is dangerous when it comes to practical decision making, since there is such heterogeneity in data reporting techniques and criteria. There is a desperate need for standard definitions of response and standardized criteria of what will be required in published results. The six papers briefly analyzed in this paper are all examples of high-quality clinical research published in a major journal. Their impact on current practice is blunted by their lack of comparability in reporting their methodology. Facile assumptions of comparability can lead to mistaken decisions about how to treat future patients and how to design future clinical research protocols. Attempts have been made in breast cancer to develop response criteria for advanced breast cancer studies [1, 3]. Major investigators and funding agencies should come together and agree on the utilization of these criteria or some acceptable modification. Once this is accomplished the editors of major journals should develop baseline criteria of what should be included in the literature reporting of a study. If general criteria become widely accepted then methodology sections can be shortened by just stating these criteria, with reference to an appropriate publication. What should no longer be accepted is the anarchy of a free-enterprise approach to data reporting in the oncologic literature.

References

1. Breast Cancer Task Force Treatment Committee, National Cancer Institute: Breast cancer: Suggested protocol guidelines for combination chemotherapy trials and for combined modality trials. Publication No. (NIH) 77-1192. Washington, D.C.: United States Department of HEW 1977
2. Bull, J. M., Tormey, D. G., Shou-Hua, L., et al.: A randomized comparative trial of adriamycin versus methotrexate in combination drug therapy. Cancer **41,** 1649–1651 (1978)
3. Hayward, J. L., Carbone, P. P., Heuson, J. G., et al.: Assessment of response to therapy in advanced breast cancer. Cancer **39,** 1289–1294 (1977)
4. Kennealey, G. T., Boston, B., Mitchell, M. S., et al.: Combination chemotherapy for advanced breast cancer: Two regimens containing adriamycin. Cancer **42,** 27–33 (1978)
5. Nemoto, T., Rosner, D., Diaz, R., et al.: Combination chemotherapy for metastatic breast cancer. Cancer **41,** 2073–2077 (1978)
6. Muss, H. B., White, D. R., Richards, F. R., et al.: Adriamycin versus methotrexate in five-drug combination chemotherapy for advanced breast cancer. Cancer **42,** 2141–2148 (1978)
7. Rubens, R. D., Begent, R. H., Knight, R. K., et al.: Combined cytotoxic and progestogen therapy for advanced breast cancer. Cancer **42,** 1680–1686 (1978)
8. Tranum, B., Hoogstraten, B., Kennedy, A., et al.: Adriamycin in combination for the treatment of breast cancer. Cancer **41,** 2078–2083 (1978)

Received March 1979/Accepted April 4, 1979