BREAST CANCER- EXPERIMENTAL AND CLINICAL ASPECTS

Editors:
H. T. Mouridsen
T. Palshof

EUROPEAN ORGANIZATION FOR RESEARCH ON TREATMENT OF CANCER (EORTC)

BREAST CANCER

Experimental and Clinical Aspects

Proceedings of the Second EORTC Breast Cancer Working Conference held in Copenhagen, 30 May to 2 June 1979

Guest Editors

T. MOURIDSEN and T. PALSHOF

Finseninstitutet, Copenhagen

Published as a supplement to the
European Journal of Cancer

PERGAMON PRESS

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon of Canada, Suite 104, 150 Consumers Road, Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, 6242 Kronberg-Taunus, Hammerweg 6, Federal Republic of Germany

First edition 1980

British Library Cataloguing in Publication Data

EORTC Breast Cancer Working Conference, *2nd, Copenhagen, 1979*
Breast cancer
1. Breast - Cancer - Congresses
I. Title II. Mouridsen, H T
III. Palshof, T IV. European Organization for Research on Treatment of Cancer
616.9'94'49 RC280.B8 79-41496

ISBN 0-08-025886-7

Published as Supplement No. 1, 1980, to the *European Journal of Cancer*

In order to make this volume available as economically and as rapidly as possible the authors' typescripts have been reproduced in their original forms. This method has its typographical limitations but it is hoped that they in no way distract the reader.

Printed in Great Britain by A. Wheaton & Co. Ltd, Exeter

Introduction

The Second E.O.R.T.C. Breast Cancer Working Conference was held in Copenhagen, May 30 to June 2, 1979.

The conference assembled 375 participants and the subjects which were discussed included aspects concerning statistical planning of trials, diagnostic methods in early and advanced breast cancer, hormone receptors, local and systemic treatment of primary and advanced disease, cell kinetics and psychological and rehabilitation aspects.

This supplement to the *European Journal of Cancer* contains the invited lectures and the free communications presented at the conference.

As reported in these papers progress in the treatment of primary and of advanced disease continues. Among the subjects which in the near future require special efforts are screening methodology and methods of selecting patients for the specific treatment modalities.

The importance of cooperation on an international basis was reemphasized at this conference both in order to ensure rapid arrangement of new progress and to ensure the validity of the conclusions of large cooperative trials.

We are very greateful to the sponsors of the conference, ICI, Pharmaceuticals Division, Danish Cancer Society, Danish Medical Council, Ministry of Education, Handelsbanken and the Finsen Institute.

July 1979

H. T. Mouridsen and T. Palshof
Guest Editors

Contents

Clinical Trials

Exclusions from Clinical Trials

K. West Andersen

The National Health Service, Store Kongensgade 1, 1264 Copenhagen K, Denmark

Correspondence to K. West Andersen

Abstract—*It is well-known that in comparing the effect of two treatments it is necessary that the two groups of patients in the trial are "equal". Likewise the results of a clinical trial are not valid for all patients but only for patients "equal" to those included in the trial. To prevent misuse of the results of a trial patients not included should be characterized in relation to all patients and for the purpose one should record patients not fulfilling the entrance criteria and patients fulfilling the entrance criterie, but not included in the trial.*

In most clinical trials some patients fulfilling the entrance criteria must be excluded e.g. some doctors may consider one of the treatments in the trial less eligible to low-risk patients and since the randomization has to be respected, such patients cannot be included in the trial. One would get a biased comparison of the treatments in the trial, if e.g. low-risk patients were withdrawn from one of the treatment groups. This means that all such exclusions have to be done before randomization. It is sometimes said that all exclusions which take place before randomization are acceptable (1). This is in some sense correct: a patient who is not included in a clinical trial cannot cause any bias in the comparison of the treatments, but observe that a patient who is randomized to one treatment but does not get that treatment may cause bias in the comparison of the treatments.

Although exclusions from a clinical trial before randomization cannot bias the comparison of the treatments, exclusions before randomization may cause the trial to be of less value. The purpose of a clinical trial is to determine the better treatment and to use that treatment in the future. Even if the better treatment is determined by a proper clinical trial one cannot use that treatment for future patients without knowing something about the excluded patients. It is of very little value to determine the better treatment for a group of patients if one cannot describe these patients.

A hypothetical example may show the bias caused by exclusions before randomization. There are, say, 200 patients who fulfill the entrance criteria in a clinical trial. For some reason a number, say 100, of these patients are excluded before randomization. The reason may be that the patient will not participate in the trial or the patient is not considered eligible although the entrance criteria is fulfilled or some mistake has occurred. The trial is then carried out for the remaining 100 patients and treatment A and B are compared. The result of the trial may be as shown in Table 1.

Table 1. Hypothetical example of results to treatment A and B in selected patients.

Treatment	No. of patients	No. of cured patients
A	50	20 = 40%
B	50	10 = 20%

The result of the trial would probably imply that all patients fulfilling the entrance criteria will be treated by A, and one would forget that half of the patients fulfilling the entrance criteria were not included in the trial. And that would be very easy to forget as probably that fact will not be mentioned in the report of the trial. When applying treatment A to all patients fulfilling the entrance criteria for the original trial one could get the result shown in Table 2.

Table 2. Hypothetical example of results to treatment A in selected and unselected patients.

Patient group	Treatment	No. of patients	No. of cured patients
"Included"	A	100	40 = 40%
"Excluded"	A	100	10 = 10%
Total	A	200	50 = 25%

Patients in the "included" group are patients "equal" to those who were included in the original trial and "excluded" patients are patients "equal" to those excluded from

the original trial although they fulfilled the entrance criteria, e.g. patients who would not participate in a clinical trial, if they were asked, or patients for whom treatment A (or B!) would be considered less eligible so that they could not be included in the original trial. For these patients who are "equal" to the excluded patients the effect of treatment A is not necessarily the same as for patients who were included in the original trial.

Another example where exclusions may cause serious bias is a multicenter trial where a standardized operation is followed by adjuvant chemotherapy. If in case of a slight deviation from the standardized operation the patient is neither included in the trial nor reported to a data collecting centre, one might, without knowing, perform a trial in which only highly selected patients are included (and the selection mechanism may be the same at all the participating hospitals). The report from the trial may recommend the standardized operation followed by the better adjuvant chemotherapy and that may cause damage if one is not aware of the fact that the patients for whom the operation were not completely successful are excluded from the trial.

When applying the result from a clinical trial one must have a patient group "equal" to those who were included in the trial. In order to make it possible to apply the results of a clinical trial in a proper way, the report from the trial should tell how many patients fulfilling the entrance criteria were not included in the trial and the reason why.

When performing a clinical trial it would therefore be of great importance to record all patients suffering from the disease and patients not included in the trial should be specially recorded, and the reason for not being included should be stated for every single patient (1,2).

The fact that the (good) results of many clinical trials are not achieved when applying the better treatment of the trial may be due to a difference between patients included in the clinical trial and patients for whom the result of the trial are applied. In order to make the results of clinical trials more useful one could propose that much effort should be done to characterize, quantitatively and qualitatively, the patients included in the trial in relation to all patients.

Another way to increase the usefulness of a clinical trial might be to include a treatment which is part of another clinical trial since that would make it possible to make a comparison of the results of the trials whatever the results of the two treatments are "equal" or not.

REFERENCES

1. R. Peto, M. C. Pike, P. Armitage, N. E. Breslow, D. R. Cox, S. V. Howard, N. Mantel, K. McPherson, J. Peto, and P. G. Smith, Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient. *Br. J. Cancer*, 34, 585 (1976).
2. Clinical trials of the treatment of breast cancer in Britain and Ireland, *Br. Med. J.* 1, 361 (1977).

*On the Analysis of Response Rates in Studies of Advanced Disease**

R. Sylvester

E.O.R.T.C. Data Center, Institut Jules Bordet, 1 rue Héger-Bordet, 1000 Brussels, Belgium

Correspondence to R. Sylvester

Abstract—*Instead of simply comparing the percentage of responders in each treatment group in studies of advanced disease, this paper advocates the use of a well known statistic test which takes the ordering of all the response categories into account.*

INTRODUCTION

In analyzing the results of studies in advanced patients with measurable disease, it is common to group evaluable patients into different categories according to the degree of tumor response measured after the start of treatment. Hayward et al (1) have for example set forth criteria for the evaluation of treatment response in advanced breast cancer patients. For the purposes of this paper we shall assume that the response to treatment falls into one of the four following categories: complete remission (C.R.), partial remission (P.R.), no change (N.C.) or progression (Prog). Although results are often reported in this manner, it is customary to analyze the data from randomized trials by comparing only the percentage of responders (C.R. or P.R.) in each treatment group. In doing so all of the available information is not used since one ignores the distinction between complete and partial remission and between no change and progression. Important differences may be missed if the treatment differences depend on the inherent ordering of the response categories which reflect the degree of tumor change. In the next section a statistic test is presented which takes this ordering into account.

METHODS

Table 1 presents the results of a hypothetical study comparing the response rates of two treatments A and B in patients with advanced breast cancer. If one compares the percentage of responders (C.R. or P.R.) in each treatment group (45/75 = 60% on treatment A and 34/75 = 45% on treatment B) using the standard chi-square test with a continuity correction, it is found that the difference is not statistically significant (P = .10).

As stated previously the above analysis does not use all the available information. Using all four response categories, one can compute within each response category the percentage of patients who receive treatment A. These percentages are given in Table 1 for the example considered (63%, 55%, 48%, 38%). If there is no difference between the treatments, these proportions should differ from one another only due to random variation. The overall test for the equality of the four proportions is in fact not significant with P = .22. This last test does not however take into consideration the ordering of the response categories and lacks the power to detect specific deviations from the hypothesis of no treatment difference.

If the ordering of the categories is now taken into consideration, one would expect the percentage of patients receiving treatment A in each response category to increase as one goes from Prog to N.C. to P.R. to C.R. in that order if in fact treatment A is better than treatment B. One way to test this hypothesis is to assign a score (1, 2, 3, 4 for example) to each response category and then compute the linear regression of the percentage of patients receiving treatment A in each response category on the score in order to determine if there is a linear trend in the proportions as one goes across the table from C.R. to Prog. The overall chi-square statistic previously computed can now be broken down into two additive components, a chi-square which tests for linear trend and a chi-square which tests for departures from linear trend. In the example given, if one assigns the scores 4 for C.R., 3 for P.R., 2 for N.C., and 1 for Prog, the test for linear trend is significant (P = .04). This indicates that the percentage of patients receiving treatment A in each response category increases as one goes from Prog to N.C. to P.R. to C.R. It can be shown that the test for trend is equivalent to testing

*This work was supported by Grant Number 2R10 CA11488-10 awarded by the National Cancer Institute, DHEW.

Table 1. Response to treatment.

Treatment	C.R.		P.R.		N.C.		Prog		Total
A	12	63%	33	55%	15	48%	15	38%	75
B	7	37%	27	45%	16	52%	25	62%	75
Total	19	100%	60	100%	31	100%	40	100%	150

Comparison of treatments A and B	*P value*
Percent C.R. or P.R. (A: 60%, B: 45%)	.10
Overall (63%, 55%, 48%, 38%)	.22
Trend (63%, 55%, 48%, 38%)	.04

whether the average score on treatment A is equal to the average score on treatment B (2). Thus a significant test for trend can be interpreted as indicating that the average response on treatment A is higher than the average response on treatment B.

While the choice of a particular set of scores may be subjective and somewhat arbitrary, one has some leeway in choosing the set of scores to be used. If for example all response categories are considered a priori to be of equal importance then the scores should be chosen to be equally spaced and any set of equally spaced scores will give the same significance level for the test for trend. Examples of such scores might be 1, 2, 3, 4; -3, -1, 1, 3; or 7, 4, 1, -2 for example. Unless a priori one wishes to emphasize a particular response category or set of response categories, the scores should be chosen to be equally spaced.

The (uncorrected) chi-square test for the comparison of the percent responders (C.R. or P.R.) in each treatment group is just a special case of the test for linear trend where now the C.R. and P.R. categories are assigned one score and the N.C. and Prog categories are assigned another score. In practice, however, it is preferable to use the continuity corrected chi-square test when two proportions are being compared. The scores may be similarly modified if a priori one wishes to test other hypotheses.

DISCUSSION

In doing a test for trend one takes into consideration the ordering of the response categories. The test for trend is more powerful than the overall test for the detection of treatment differences if a linear trend is present, which may well be the case if the two treatments differ in efficacy. The decision however concerning choice of scores to be used should be made prior to the start of the study so that the choice of the hypothesis to be tested does not depend on the results of the study.

The formulas used in the above calculations can be found in the appendix which follows.

APPENDIX

The notation used in the appendix follows that of Armitage (3). Suppose that you have k response categories for each of two treatments with a score X assigned to each category. Then a 2 × k contingency table of the treatment results can be constructed as shown in Table 2.

Where X_i = the score associated with response category i
r_i = the number of patients in response category i receiving treatment A
$n_i - r_i$ = the number of patients in response category i receiving treatment B
n_i = the total number of patients in response category i
$P_i = r_i/n_i$ = the proportion of patients in response category i receiving treatment A
$R = \Sigma r_i$ = total number of patients receiving treatment A
$N-R = \Sigma\, n_i - r_i$ = total number of patients receiving treatment B
$N = \Sigma n_i$ = total number of patients
$P = R/N$ = overall proportion of patients receiving treatment A

where all summations are from i = 1 to k. Then for k > 2

$$X^2_{k-1} = \frac{\Sigma(r_i^2/n_i) - R^2/N}{P(1-P)}$$

provides an overall test for the equality of the proportions P_i. Under the null hypothesis of no treatment difference, X^2_{k-1} is approximately distributed as chi-square with k-1 degrees of freedom. If one wishes to ask whether there is a significant trend in the proportions P_i from response category 1 to response category k (a trend of P_i with x_i) then the statistic X^2_{k-1} can be broken down into two additive components:

(1) a test for linear trend

$$X^2_1 = \frac{N(N\Sigma r_i X_i - R\Sigma n_i X_i)^2}{R(N-R)(N\Sigma n_i X_i^2 - (\Sigma n_i X_i)^2)}$$

which under the null hypothesis of no trend is distributed approximately as chi-square with 1 degree of freedom and

Table 2.

Group	1	2	...	i	...	k	Total
Score	X_1	X_2	...	X_i	...	X_k	
Treatment A	r_1	r_2	...	r_i	...	r_k	R
Treatment B	$n_1 - r_1$	$n_2 - r_2$	...	$n_i - r_i$	...	$n_k - r_k$	N-R
Total	n_1	n_2	...	n_i	...	n_k	N
Proportion Treatment A	P_1	P_2		P_i		P_k	$P = \frac{R}{N}$

(2) a test for departure from linear trend

$$X^2_{k-2} = X^2_{k-1} - X^2_1$$

which under the null hypothesis of no departure from linear trend is approximately distributed as chi-square with k-2 degrees of freedom.

For the case k = 2 the test for trend is equivalent to the uncorrected chi-square statistic for the comparison of two percentages. However, in this case it is preferable to use the following continuity corrected statistic:

$$X^2_1 = \frac{(|r_1(n_2-r_2)-r_2(n_1-r_1)| - N/2)^2 N}{R(N-R)n_1n_2}$$

which under the null hypothesis of no treatment differences is asymptotically distributed as chi-square with 1 degree of freedom. For further details please consult Armitage (3,4), Cochran (2), Everitt (5), or Fleiss (6).

REFERENCES

1. J. L. Hayward, P. P. Carbone, J. C. Heuson, S. Kumaoka, A. Segaloff and R. D. Rubens, Assessment of response to therapy in advanced breast cancer. *Cancer* 39, 1289 (1977).
2. W. G. Cochran, Some methods for strengthening the common χ^2 tests. *Biometrics* 10, 417 (1954).
3. P. Armitage, *Statistical methods in medical research*, Blackwell Scientific Publications, Oxford (1971).
4. P. Armitage, Tests for linear trends in proportions and frequencies: *Biometrics* 11, 375 (1955).
5. B. S. Everitt, *The analyses of contingency tables*, Chapman and Hall Ltd., London (1977).
6. J. L. Fleiss, *Statistical methods for rates and proportions*, John Wiley and Sons, New York (1973).

Diagnostic Methods in Early and Late Breast Cancer