Adjuvant Chemotherapy for Breast Cancer

A Pooled Estimate Based on Published Randomized Control Trials

Harvey N. Himel, MD, MPH; Alessandro Liberati, MD; Richard D. Gelber, PhD; Thomas C. Chalmers, MD

The use of adjuvant chemotherapy for treating patients with operable breast cancer remains a worldwide controversy. Using the data from published randomized control trials with a minimum two-year follow-up, pooled estimates of relapse-free survival rates and overall survival rates were calculated. Relapse-free survival rates were improved by 12.5% (95% confidence interval [CI] ±4.5%) at three years and by 8% (CI ±6%) at five years, with studies using multiple agents showing a greater effect. A significant advantage was also present in overall survival rates at three years, but only for studies involving multiple agents (4%±3.5%). Results from combining data for other types of trials were inconclusive. The use of this method is presented to illustrate its value as an explicit and systematic one for combining data from several randomized control trials in assessing a therapeutic controversy.

(*JAMA* 1986;256:1148-1159)

DESPITE 30 years of research and development, the role of chemotherapy in the adjuvant treatment of primary breast cancer remains controversial.[1-8] During that time, at least 12 different drugs have been tested alone or in combination in at least 31 published randomized control trials (RCTs). Although almost 10 000 patients have been enrolled in these trials, the degree to which chemotherapy improves the overall health of a patient with primary breast cancer and increases her life expectancy remains undecided. Nevertheless, every practitioner must decide with each such patient what is the best course of treatment based on the available evidence.

According to one recent survey, most American physicians now consider some form of adjuvant chemotherapy to be appropriate treatment for women with stage II breast cancer.[9] Traditional reviews of the literature have tended to promote the use of chemotherapy[2-5] although there have been some dissenters.[6-8] These reviews, however thorough, derive conclusions from qualitative reviews of the various studies' results and the disparate conclusions of the individual investigators.

This article presents an exhaustive compilation of available data and applies a quantitative method to combine the data presented in RCTs into a weighted mean.[10,11] The method is offered as an approach to resolving systematically a therapeutic controversy by pooling data available from published RCTs. Furthermore, the specific results obtained are offered as the basis for quantitative policy formation through such methods as cost-effectiveness analysis, as well as for more precise prognostication by clinicians and patients.

The validity of pooling similar RCTs has been challenged[12] because of three major causes of heterogeneity in the data to be combined: differences in patients to be studied, differences in therapeutic regimens applied, and differences in the quality of the RCTs. While the first two can be corrected for to some extent by analytic methods such as stratified analysis, the relative effect of bias on the validity of pooling has not been assessed previously. In another publication, we analyzed the quality of the RCTs for adjuvant chemotherapy of stage II breast cancer,[13] using a method previously described[14] that expressly evaluates several characteristics of a study that are associated with observer bias. We have incorporated this information into this data pooling report.

A wide variation in the quality of the research design and execution and grave deficiencies in the reporting of essential clinical data were found in the studies read.[13] However poor the quality of the published literature, it is all that is directly available to the practicing physician.

The combined data suggest that the less toxic single-drug regimens delay relapse in a small percentage of patients, but do not influence overall survival. However, the more toxic multiple-drug regimens postpone relapse for a larger percentage of patients and improve overall survival, especially in certain subgroups. Nevertheless, better data must be gathered over longer periods of time to be confident about the place of chemotherapy for all groups of patients.

METHODS

We examined 31 RCTs reported in articles written in English over the last 30 years and published through Decem-

From the Departments of Health Policy & Management (Drs Himel and Chalmers), Epidemiology (Dr Liberati), and Biostatistics (Dr Gelber), Harvard School of Public Health, Boston; the Department of Biostatistics, Dana-Farber Cancer Institute (Dr Gelber), Boston; and the Clinical Trials Unit, Mount Sinai School of Medicine, New York (Dr Chalmers).

Reprint requests to Mount Sinai School of Medicine, 1 Gustave L. Levy Place, New York, NY 10029 (Dr Chalmers).

1. SHORT-TERM CHEMOTHERAPY

A. Studies With Multiple Reports

Set 1

1a. Noer RJ (Surgical Adjuvant Chemotherapy Breast Group): Adjuvant chemotherapy: Thio-Tepa with radical mastectomy in the treatment of breast cancer. *Am J Surg* 1963;106:405-412.

2a. Fisher B, Ravdin RG, Ausman RK, et al: Surgical adjuvant chemotherapy in cancer of the breast: Results of a decade of cooperative investigation. *Ann Surg* 1968;168:337-356.

3a. Fisher B, Slack N, Katrych D, et al: Ten year follow-up results of patients with carcinoma of the breast in a co-operative clinical trial evaluating surgical adjuvant chemotherapy. *Surg Gynecol Obstet* 1975;140:528-534.

4a. Surgical Adjuvant Chemotherapy Breast Group: Breast adjuvant chemotherapy: Effectiveness of Thio-Tepa (Triethylenethiophosphoramide) as adjuvant to radical masectomy for breast cancer. *Ann Surg* 1961;54:629-647.

Set 2

5a. Nissen-Meyer R, Kjellgren K, Malmio K, et al: Surgical adjuvant chemotherapy: Results with one short course with cyclophosphamide after mastectomy for breast cancer. *Cancer* 1978;1:2088-2096.

6a. Nissen-Meyer R, Kjellgren K, Mansson B: Adjuvant chemotherapy in breast cancer. *Recent Results Cancer Res* 1982;80:142-148.

B. Study With a Single Report

7a. Finney R: Adjuvant chemotherapy in the radical treatment of carcinoma of the breast: A clinical trial. *AJR* 1971;111:137-141.

2. LONG-TERM CHEMOTHERAPY WITH UNTREATED CONTROL GROUPS

A. Studies With Multiple Reports

Set 1

8a. Bonadonna G, Brusamolino E, Valagussa P, et al: Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* 1976;294:405-410.

9a. Bonadonna G, Rossi A, Valagussa P, et al: The CMF program for operable breast cancer with positive axillary nodes. *Cancer* 1977;39(suppl 6):2904-2915.

10a. Rossi A, Bonadonna G, Tancini G, et al: Trials of adjuvant chemotherapy in breast cancer: The experience of the Instituto Nazionale Tumori of Milan. *Eur J Cancer* 1980;(suppl 1):149-156.

11a. Rossi A, Bonadonna G, Valagussa P, et al: Multimodal treatment in operable breast cancer: Five-year results of the CMF programme. *Br Med J* 1981;282:1427-1431.

12a. Bonadonna G, Rossi A, Tancini G, et al: CMF adjuvant programs at the Milan Cancer Institute, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 66-73.

Set 2

13a. Fisher B, Carbone P, Economou SG, et al: L-Phenylalanine mustard (L-Pam) in the management of primary breast cancer: A report of early findings. *N Engl J Med* 1975;292:117-122.

14a. Fisher B, Glass A, Redmond C, et al: L-Phenylalanine mustard (L-Pam) in the management of primary breast cancer: An update of earlier findings and a comparison with those utilizing L-Pam plus 5-fluorouracil (5-FU). *Cancer* 1977;39(suppl 6):2883-2903.

15a. Fisher B, Redmond C, Fisher ER, et al: The contribution of recent NSABP clinical trials of primary breast cancer therapy to an understanding of tumor biology: An overview of findings. *Cancer* 1980;46(suppl):1009-1025.

16a. Fisher B, Redmond C, Wolmark N, et al: Disease-free survival at intervals during and following completion of adjuvant chemotherapy: The NSABP experience from three breast cancer protocols. *Cancer* 1981;48:1273-1280.

17a. Fisher B, Fisher ER, Redmond C, et al: A brief overview from NSABP trials of adjuvant therapy, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 55-65.

18a. Wolmark N, Fisher B: Adjuvant chemotherapy in stage II breast cancer: A brief overview of the NSABP clinical trials. *World J Surg* 1985;9:699-706.

Set 3

19a. Multicentre Breast Cancer Chemotherapy Group: Multimodal therapy for histological stage-II breast cancer. *Lancet* 1977;2:396-397.

20a. Wheeler TK, Edelstyn GA, Bates TS, et al: Adjuvant chemotherapy with four drugs for stage II breast cancer. *Eur J Cancer* 1980;(suppl 1):161-163.

Set 4

21a. Senn HJ, Jungi WF, Amgwerd R: Adjuvant chemoimmunotherapy with LMF + BCG in node-negative and node-positive breast cancer patients: Intermediate report of a randomized trial in patients. *Antibiot Chemother* 1978;24:213-228.

22a. Jungi WF, Senn HJ, Amgwerd R, et al: Divergent effect of adjuvant chemo-imunotherapy on recurrence rates in node negative and node-positive breast cancer patients. *Eur J Cancer* 1980;(suppl 1):169-172.

23a. Senn H, Jungi WF, Amgwerd R: Chemoimmunotherapy with LMF and BCG in node negative and node-positive breast cancer, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer III.* New York, Grune & Stratton, 1981, pp 385-393.

24a. Senn HJ, Amgwerd R, Jungi WF, et al: Adjuvant chemo-immunotherapy with LMF plus BCG in node-negative and node-positive breast cancer: Intermediate report at 4 years. *Recent Results Cancer Res* 1982;80:177-184.

25a. Senn HJ, Jungi WF, Amgwerd R, et al: Adjuvant chemoimmuotherapy with LMF + BCG in node-negative and node-positive breast cancer: Eight-year results, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 90-101.

Set 5

26a. Long RT, Donegan WL, Evans AM: Extended surgical adjuvant chemotherapy for breast carcinoma. *Am J Surg* 1969;117:701-704.

27a. Donegan WL: Extended surgical adjuvant thiotepa for mammary carcinoma. *Arch Surg* 1974;109:187-192.

28a. Kardinal CG, Donegan WL: Second cancers after prolonged adjuvant thiotepa for operable carcinoma of the breast. *Cancer* 1980;45:2042-2046.

Set 6

29a. Morrison JM, Howell A, Grieve RJ, et al: The West Midlands Oncology Association Trials of adjuvant chemotherapy for operable breast cancer, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer III.* New York, Grune & Stratton, 1981, pp 403-410.

30a. Morrison JM, Howell A, Grieve RJ, et al: The West Midlands Oncology Association trials of adjuvant chemotherapy for operative breast cancer, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer IV.* New York, Grune & Stratton, 1984, pp 253-259.

B. Studies With Single Reports

31a. Leiberman DP, Berstock DA, Houghton J, et al: Oral adjuvant therapy in breast carcinoma: A multicentre trial. *Cancer Treat Rev* 1979;6(suppl):91-96.

32a. Kaufmann M, Fournier DV, Sievers H, et al: Adjuvant chemotherapy with chlorambucil and 5-fluorouracil in primary breast cancer (Cooperative Study Heidelberg). *Eur J Cancer* 1980;(suppl):157-160.

33a. Koyama H, Wada T, Takahashi Y, et al: Surgical adjuvant chemotherapy with mitomycin C and cyclophosphamide in Japanese patients with breast cancer. *Cancer* 1980;46:2373-2379.

34a. Rubens RD, Knight RK, Fentiman IS, et al: Controlled trial of adjuvant chemotherapy with melphalan for breast cancer. *Lancet* 1983;1:839-843.

35a. Ludwig Breast Cancer Study Group: Randomized trial of chemo-endocrine therapy, endocrine therapy and mastectomy alone in post-menopausal patients with operable breast cancer and axillary-node metastasis. *Lancet* 1984;1:1256-1260.

36a. Howell A, Bush H, George WD, et al: Control trial of adjuvant chemotherapy with cyclophosphamide methotrexate and 5-fluorouracil for breast cancer. *Lancet* 1984;2:307-311.

37a. Smith DC, Crawford D, Dykes EH, et al: Adjuvant radiotherapy and chemotherapy in breast cancer, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer IV.* New York, Grune & Stratton, 1984, pp 283-289.

38a. Tormey DC, Taylor SG, Gray R, et al: Postmenopausal, node-positive comparison of observation with CMFP and CMFP + tamoxifen adjuvant therapy: An Eastern Cooperative Oncology Group trial, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 110-116.

3. LONG-TERM CHEMOTHERAPY WITH TREATED CONTROL GROUPS

A. Studies With Multiple Reports

Set 1

39a. Tancini G, Bajetta E, Marchini S, et al: Preliminary 3-year results of 12 vs. 6 cycles of surgical adjuvant CMF in premenopausal breast cancer. *Cancer Clin Trials* 1979;2:285-292.

40a. Tancini G, Bonadonna G, Valgussa P, et al: Adjuvant CMF in breast cancer: Comparative five year results of 12 versus six cycles. *J Clin Oncol* 1983;1:2-10.

Set 2

41a. Andersen KW, Mouridsen HT, Castbert T, et al: Organization of the Danish adjuvant trials in breast cancer. *Dan Med Bull* 1981; 28:102-106.

42a. Brincker H, Mouridsen HT, Andersen KW, et al: Adjuvant chemo-

(Continued on p 1150.)

Appendix 1.—List of Trial Reports Read (cont)

therapy with cyclophosphamide or CMF in premenopausal women with stage II breast cancer. *Breast Cancer Res Treat* 1983; 3:91-95.

43a. Mouridsen HT, Rose C, Brincker H, et al: Adjuvant systemic therapy in high risk breast cancer: The Danish Breast Cooperative Group's trials of cyclophosphamide or CMF in premenopausal and tamoxifen in postmenopausal patients, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 117-128.

Set 3

44a. Weiss RB, Tormey DC, Holland F, et al: A randomized trial of post-operative five- versus three-drug chemotherapy after masectomy: A cancer and Leukemia Group B (CALGB) study. *Recent Results Cancer Res* 1982;80:170-176.

45a. Tormey DC, Weinberg VE, Holland JF, et al: A randomized trial of five and three drug chemotherapy and chemoimmunotherapy in women with operable node positive breast cancer. *J Clin Oncol* 1983;1:138-145.

46a. Tormey DC: Adjuvant systemic therapy in postoperative node-positive patients with breast carcinoma: The CALGB trial and ECOG premenopausal trial, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 155-165.

Set 4

47a. Caprini JA, Oviedo MA, Cuningham MP, et al: Adjuvant chemotherapy for stage II and III breast carcinoma. *JAMA* 1980;244:243-246.

48a. Cohen E, Scanlon EF, Caprini JA, et al: Follow-up adjuvant chemotherapy and chemoimmunotherapy for stage II and III carcinoma of the breast. *Cancer* 1982;49:1754-1761.

Set 5

49a. Glucksberg H, Rivkin SE, Rasmussen S: Combination chemotherapy (CMFVp) versus L-phenylalanine mustard (L-PAM) for operable breast cancer with positive axillary nodes: A Southwest Oncology Study. *Cancer* 1982;50:423-434.

50a. Knight WA, Rivkin SE, Glucksberg H, et al: Adjuvant therapy of breast cancer: The Southwest Oncology Group experience. *Breast Cancer Res Treat* 1983;3(suppl 1):27-33.

51a. Rivkin SE, Glucksberg H, Foulkes M: Adjuvant therapy of breast cancer: A Southwest Oncology Group experience, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 166-174.

Set 6

52a. Ahmann DL, Scanlon PW, Bisel HF, et al: Repeated adjuvant chemotherapy with phenylalanine mustard or 5-fluorouracil, cyclophosphamide, and prednisone with or without radiation, after mastectomy for breast cancer. *Lancet* 1978;1:893-896.

53a. Ahmann DL, O'Fallon JR, Scanlon PW, et al: A preliminary assessment of factors associated with recurrent disease in a surgical adjuvant clinical trial for patients with breast cancer with special emphasis on the aggressiveness of therapy. *Am J Clin Oncol* 1982;5:371-381.

Set 7

54a. Jungi WF, Alberto P, Brunner KW, et al: Short- or long-term adjuvant chemotherapy for breast cancer, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer III.* New York, Grune & Stratton, 1981, pp 395-402.

55a. Jungi WF, Alberto P, Brunner KW, et al: Short- or long-term chemotherapy for node-positive breast cancer: LMF six versus 18 cycles: SAKK study 27/76, in Senn HJ (ed): *Adjuvant Chemotherapy of Breast Cancer,* vol 96, in *Recent Results in Cancer Research.* Berlin, Springer-Verlag, 1984, pp 175-177.

Set 8

56a. Velez-Garcia E, Moore M, Vogel CL, et al: Postmastectomy adjuvant chemotherapy with or without radiation therapy in women with operable breast cancer and positive axillary lymph nodes: The Southeastern Cancer Group experience. *Breast Cancer Res Treat* 1983;3(suppl 1):49-60.

57a. Velez-Garcia E, Moore M, Vogel CL, et al: Postsurgical adjuvant chemotherapy with or without radiation in women with breast cancer and positive axillary nodes: The Southeastern Cancer Study Group (SECSG) experience, in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer IV.* New York, Grune & Stratton, 1984, pp 273-282.

B. Studies With Single Reports

58a. Cooper MR, Rhyme AL, Muss HB, et al: A randomized comparative trial of chemotherapy and irradiation therapy for stage II breast cancer. *Cancer* 1981;47:2833-2839.

59a. Chlebowski RT, Weiner JM, Luce J, et al: Significance of relapse after adjuvant treatment with combination chemotherapy or 5-fluorouracil alone in high-risk breast cancer: A Western Cancer Study Group project. *Cancer Res* 1981;41:4399-4403.

60a. Carpenter JT, Maddox WA, Laws HL, et al: Favourable factors in the adjuvant therapy of breast cancer. *Cancer* 1982;50:18-23.

61a. Misset JL, DeVassal F, Jasmis C, et al: Five year results of the French adjuvant trial for breast cancer comparing CMF to a combination of Adriamycin (ADM), vincristine (VCR), cyclophosphamide (CPM), and 5-fluorouracil (5FU), in Salmon SE, Jones SE (eds): *Adjuvant Therapy of Cancer IV.* New York, Grune & Stratton, 1984, pp 243-251.

ber 1984. Articles were retrieved by scanning Current Contents and Medical Subject Headings' (MeSH) key words (viz, neoplasm, human, breast disease, and random allocation), by inspecting the bibliographies of original and review articles on the treatment of breast cancer, and by direct inquiry with the principal investigators of articles found through one of the previous methods. Moreover, to complete our search and estimate the number of unpublished studies, we contacted all principal investigators of RCTs listed in the National Cancer Institute file of closed and active trials (CLINPROT). For the purpose of this study, we considered only results that were fully reported in journal articles or monographs and did not include any data from abstracts or presentations given at meetings. Information reported in letters to the editor was considered only insofar as it might update results of a study that had already been extensively reported in an earlier published article. We made one exception to the cutoff date of December 1984, to include data from the NSABP B-05 trial (L-phenylalanine mustard vs placebo) for pooling (Appendix I, 18a).

Studies were considered eligible for data pooling if (1) the authors stated that they had assigned patients to treatment on a random basis, (2) the treatment group differed from the control group with regard to regimen of chemotherapy received, and (3) all patients in the study had been followed up for at least two years after their enrollment into the study.

Treatment effects were assessed in terms of relapse-free survival (RFS), as defined in the individual studies, and overall survival (OS), expressed as percentages at three, five, and (in the case of two trials of short-term treatment) ten years of follow-up.

Data were gathered from manuscripts (Appendix 1) using the most recent report. Estimated rates for RFS and OS were obtained from numbers reported in tabular form. When these data were not available, data were extracted by applying a draftsman's T-square to the published life-table curves. The data for each individual study used to obtain pooled estimates are reported in Appendixes 2 through 4, together with the source of the information.

Most studies reported results before all patients had been followed up for five years. The sample size used for pooling the data extracted was adjusted downward in this circumstance, as follows: we considered only those patients who were known to have reached the time point being measured without relapse or death (number at risk for a given time point) and extrapolated the number of patients expected to fail based on the observed rate of failure reported in the articles read. Thus, an approximation for the effective sample size (N′) was derived according to the following formula:

$$N' = Nat/Or$$

where *Or* was the observed rate of "success" (RFS or OS) derived at the time point of interest; and *Nat*, the number of patients at risk at the time point of interest.

For example, consider a hypothetical trial with 100 patients in the control

Appendix 2.—Data Used in Short-term Trials of Chemotherapy Relapse-Free Survival

Study and Source*	5-y Data				10-y Data			
	No. Treated	Success Rate	No. of Controls	Success Rate	No. Treated	Success Rate	No. of Controls	Success Rate
All Patients								
1. Scandinavian Adjuvant Chemotherapy Group[6a] (p 144, Fig 3)	507	0.65	519	0.56	507	0.56	519	0.48
2. NSABP I (B-01)†[3a] (p 530, Table 2)	370	0.62	370	0.60	327‡	0.50	323	0.50
Premenopausal Patients								
1. Scandinavian Adjuvant Chemotherapy Group[6a] (p 146, Fig 5)	242	0.71	265	0.58	. . .	. . .	. . .	. . .
2. NSABP I (B-01)[3a] (p 530, Table 2)	98	0.61	112	0.53	. . .	. . .	. . .	. . .
Postmenopausal Patients								
1. Scandinavian Adjuvant Chemotherapy Group[6a] (p 146, Table 2)	265	0.64	254	0.54	. . .	. . .	. . .	. . .
2. NSABP I (B-01)[3a] (p 530, Table 2)	272	0.62	258	0.64	. . .	. . .	. . .	. . .
Patients With Nodal Involvement								
1. Scandinavian Adjuvant Chemotherapy Group[6a] (p 146, Fig 4)	198	0.45	218	0.35	. . .	. . .	. . .	. . .
2. NSABP I (B-01)†[3a] (p 530, Table 2)	186	0.44	172	0.35	. . .	. . .	. . .	. . .
Overall Survival: All Patients								
1. Scandinavian Adjuvant Chemotherapy Group[6a] (p 144, Fig 3)	507	0.75	519	0.68	507	0.59	519	0.52
2. NSABP I (B-01)[3a] (p 531, Table 3)	414	0.65	406	0.64	397‡	0.46	390	0.46

*All source references refer to articles listed in Appendix 1.

†NSABP indicates National Surgical Adjuvant Breast Project.

‡Sample size decreased by authors who reported original study in tabular form at five and ten years.

group, only 63 of whom were enrolled in the study more than three years ago and remain at risk. If the reported three-year OS was 90%, then an approximation to the effective sample size for the three-year time point would be 63/0.90, or 70 patients. Data were not considered for extrapolation at time points where N' was less than one third of N, the initial sample size.

At each time point (three, five, and ten years), we derived effective sample sizes for the treated and control groups (n_{ti} and n_{ci} for the i-th study), and observed rates for the event of interest for each of the two groups (r_{ti} and r_{ci} for the i-th study). The rate difference ($RD_i = r_{ti} - r_{ci}$) was used as a measure of treatment effect for the i-th study. Data from the different studies were combined using the method described by DerSimonian and Laird.[10] Studies were pooled by weights utilizing the inverse of the variance. Hence, larger studies have smaller corresponding variances and make greater contributions to the pooled estimate.

The method also provided a statistic, Q, that may be used to test for the homogeneity of treatment effects across studies and also may appropriately serve to increase the width of the confidence interval of the pooled estimate to account for differences between studies. Further details of the computational methods are presented in Appendix 5.

The 31 RCTs included in the analyses have been divided into four categories as follows (Table 1). Fourteen RCTs compared various chemotherapy regimens—used for a minimum of six months—with placebo or no further treatment following mastectomy. Three RCTs compared a short course of perioperative chemotherapy, using a single-drug regimen, with an untreated control group. Eleven RCTs compared different regimens with each other, ie, single-agent vs multiple-agent chemotherapy. Finally, three RCTs compared different durations of time for administering the same regimen of chemotherapy.

In the presentation of the data, the studies in each graph are presented in the order of a quality score assigned to each study. This quality score has been determined in a previous study, which investigated the quality of 63 RCTs on the treatment of primary breast cancer.[13] The studies were read and evaluated according to a standardized checklist that considered issues of both internal (scientific) validity and external validity (generalizability). Internal validity scores were considered most appropriate for the data pooling, as these evaluated items of design, execution, and data analysis that could influence the amount of bias in study results. Initially all available studies were pooled and then a sensitivity analysis was performed by eliminating those with the lowest internal validity quality scores.

RESULTS

The general characteristics of the 31 RCTs included in our analyses are summarized in Table 1. The median sample size was 325 patients, with a range of 62 to 1026. The trials are listed in descending order of quality score

Appendix 3.—Data Used in Combination of Long-term Trials of Chemotherapy With Untreated Controls

Study and Source*	2-3-y Data				5-y Data			
	No. Treated	Success Rate	No. of Controls	Success Rate	No. Treated	Success Rate	No. of Controls	Success Rate
All Patients								
1. Ludwig[35a] (p 1258, Fig 1)								
CMFpT	154	0.77	156	0.55	...	...	...	...
pT CMFpT	153	0.62	...	...	...	...	...	...
2. NSABP II (B-05)[18a] (p 700, Fig 1)	179	0.62	170	0.56	179	0.53	170	0.47
3. Milan I[12a] (p 67, Fig 1)	207	0.67	179	0.49	207	0.55	179	0.45
4. Guy's II (CMF)[36a] (p 308, Fig 1A)	84†	0.73	86†	0.59	...	...	...	...
5. ECOG[38a]‡ (p 112, Fig 1)								
CMFPT	69	0.72	82	0.53	...	...	...	...
CMFP	73	0.58	...	...	...	...	...	...
6. OSAKO[25a] (p 94, Fig 3A)	117	0.72	123	0.57	117	0.59	123	0.52
7. Guy's I L-Pam[34a] (p 841, Fig 1)	187	0.60	183	0.58	95†	0.56	94†	0.48
8. MBCCG[20a] (p 163, Table 4)	97†	0.62	98†	0.49	...	...	...	...
9. Ellis Fischel[27a] (p 190, Fig 2)†	55†	0.62	45†	0.56	...	...	...	...
Premenopausal Patients								
1. NSABP II (B-05)[18a] (p 700, Fig 2)	59	0.68	61	0.53	59	0.62	61	0.42
2. Milan I[12a] (p 68, Fig 3)	103	0.74	86	0.46	103	0.61	86	0.42
3. Guy's II CMF[36a] (p 308, Fig 1B)	42†	0.67	40†	0.52	...	...	...	...
4. Guy's I L-Pam[34a] (p 841, Fig 2)	79	0.64	77	0.57	45†	0.62	46†	0.52
Postmenopausal Patients								
1. Ludwig[35a] (p 1258, Fig 1)								
CMFpT	154	0.77	156	0.55	...	...	...	...
pT	153	0.62	...	...	...	...	...	...
2. NSABP II (B-05)[18a] (p 700, Fig 2)	120	0.58	109	0.59	120	0.49	109	0.49
3. Milan I[12a] (p 68, Fig 3)	104	0.62	93	0.54	104	0.50	93	0.47
4. Guy's II CMF[36a] (p 308, Fig 1C)	45†	0.73	31†	0.68	...	...	...	...
5. ECOG[38a]‡ (p 112, Fig 1)								
CMFPT	69	0.72	82	0.53	...	...	...	...
CMFP	73	0.58	...	...	...	...	...	...
6. OSAKO[25a] (p 95, Fig 4A)	60	0.74	54	0.62	...	...	...	...
7. Guy's I L-Pam[34a] (p 841, Fig 3)	108	0.58	106	0.58	48†	0.52	45†	0.47
Patients With 1-3 Nodes Involved								
1. Milan I[11a] (p 1428, Fig 3)	140	0.74	126	0.54	140	0.69	126	0.48
3. Guy's I L-Pam[34a] (p 841, Fig 4)	109	0.73	115	0.66	56†	0.71	62†	0.60
Patients With >3 Nodes Involved								
2. Milan I[11a] (p 1428, Fig 3)	67	0.53	53	0.36	67	0.41	53	0.33
3. Guy's I L-Pam[34a] (p 842, Fig 5)	78	0.43	68	0.41	36†	0.36	30†	0.27

Appendix 3.—Data Used in Combination of Long-term Trials of Chemotherapy With Untreated Controls (cont)

Study and Source*	2-3-y Data				5-y Data			
	No. Treated	Success Rate	No. of Controls	Success Rate	No. Treated	Success Rate	No. of Controls	Success Rate
1. Ludwig[35a] (p 1258, Fig 1)			Overall Survival: All Patients					
CMFpT	154	0.88	156	0.88	...	...	...	...
pT	153	0.88	...	...	...	...	...	...
2. NSABP II (B-05)[17a] (p 62, Fig 6)	179	0.78	170	0.76	179	0.67	170	0.62
3. Milan I[12a] (p 67, Fig 1)	207	0.85	179	0.79	207	0.72	179	0.67
4. Guy's II CMF[36a] (p 309, Fig 2)	165†	0.85	162†	0.78	...	...	...	...
5. ECOG[38a] (p 112, Fig 2)								
CMFPT	69	0.90	82	0.86	...	...	...	...
CMFP	73	0.86	...	...	...	...	...	...
6. OSAKO[25a] (p 94, Fig 3B)	117	0.91	123	0.85	117	0.81	123	0.72
6. Guy's I L-Pam[34a] (p 842, Fig 7)	187	0.70	183	0.83	98†	0.65	96†	0.68
7. Ellis Fischel[28a] (p 2045, Fig 1)	90	0.70	77	0.74	90	0.53	77	0.58

*All source references refer to reports listed in Appendix 1. NSABP indicates National Surgical Adjuvant Breast Project; ECOG, Eastern Cooperative Oncology Group; OSAKO, the East Switzerland Cooperative Oncology Group; MBCCG, Multicenter Breast Cancer Chemotherapy Group; C, cyclophosphamide; M, methotrexate; F, fluorouracil; p, prednisone (low dose); P, prednisone (high dose); T, tamoxifen; L-Pam, L-phenylalanine mustard.

†Sample size decreased to N′ (see text) because study not mature at this point.

‡Data from two treatment arms arithmetically averaged in pooling.

Appendix 4.—Data Used in Long-term Trial of Chemotherapy With Treated Control Group

Study and Source*	2-3-y Data			
	No. Treated	Success Rate	No. of Controls	Success Rate
SWOG[49a] (p 531, Fig 2)	166	0.76	183	0.56
	Overall Survival			
SWOG[49a] (p 531, Fig 1)	172	0.82	186	0.72

*Source reference refers to article listed in Appendix 1. SWOG indicates Southwestern Oncology Group.

within each of the four subcategories. The median quality score for all 31 RCTs was 45%, with no significant difference in median score among the four groups. Median scores were 43% for trials comparing long-term regimens with an untreated control group, 45% for trials comparing different regimens of chemotherapy, 46% for those reporting short-term chemotherapy, and 44% for those comparing different durations of the same regimen.

Figure 1 shows the observed rate differences at three and five years' follow-up for RFS, and Fig 2 shows the OS results among trials of long-term chemotherapy against an untreated control group. Data on RFS and OS were available at three years from only nine and eight of the 14 published studies, respectively. Results at five years' follow-up could be obtained for the two end points from only five and four studies, respectively. The pooled results for studies using a single-agent regimen, a multiple-agent regimen, and all studies combined are each displayed at the bottom of the figures. Of these three pooled estimates, only the combination of all trials and the subgroup of multiple drug regimen trials show a statistically significant increase in RFS both at three years and at five years. Pooled results for overall survival were not statistically significant at three years (RD ± 95% confidence interval [CI]=2% ± 3%) nor at five years (4% ± 5%), based on the analysis of data from studies that are mature to that point. However, multiple-agent chemotherapy regimens showed a consistently greater treatment effect than single-agent treatments. Pooling of five such multiple-agent studies with data available at three-year follow-up showed an estimated treatment effect on OS of 4% (95% CI, 3.5%) while at five years' follow-up, data from two studies were available (Fig 2), and that RD was statistically significant (7% ± 6.5%).

The RFS results of pooling data within patient subpopulations based on menopausal status are presented in Figs 3 and 4. Some improvement in RFS was seen for both follow-up periods with premenopausal patients and at three but not five years in postmenopausal patients. Too few studies reported data of the subgroups of one to three and four or more axillary lymph nodes involved to allow pooling, but again the multiple-agent regimen demonstrated a greater treatment in both groups for RFS at three and five years.

With studies comparing two forms of adjuvant chemotherapy, multiple-agent chemotherapy appeared to perform better than single-agent chemotherapy for RFS and OS at the three- and five-year follow-up time points, but results were not consistent. Too few data were available from studies to pool results of single- vs multiple-agent therapy.

The two studies on short-term perioperative chemotherapy compared with placebo that had published data

available for five and ten years are presented in Fig 5. The results of the two studies are contradictory, and the pooled analysis indicated an overall treatment advantage that was not statistically significant, both for RFS and OS.

Altogether, there were 30 sets of data that were pooled and presented in the accompanying figures. The Q statistic measuring study heterogeneity was never significant. In only six analyses was the magnitude of Q large enough to increase the estimated SE for the pooled results.

COMMENT

In this study we attempted a systematic combination of results from published RCTs on adjuvant chemotherapy for operable breast cancer. Partly because of the poor quality of the reporting and partly due to the fact that for some studies only preliminary results are available in fully published form, we could use only about half of the potentially available information. The combined evidence from the first group of RCTs in which patients treated for more than six months were compared with an untreated control group showed a marked effect of chemotherapy in delaying recurrence and a smaller but statistically significant effect on overall survival when using multiple-agent therapy.

No conclusion can be drawn from the examination of trials comparing treatment with a single-agent to a multiple-drug regimen; only one study had enough data to be pooled with a length of follow-up of three years. Finally, while the combination of the two studies on short-term chemotherapy does not suggest definite conclusions about the efficacy of this modality of treatment, this regimen is now the subject of new interest and investigation.[15]

A recent meeting of almost all trialists studying the impact of adjuvant treatment for breast cancer has been reported.[16] Data on over 10 000 patients randomized in trials involving the evaluation of adjuvant chemotherapy contributed to that overview. Our intent was to evaluate the evidence currently available in published form.

Poor quality of reporting of the data presented in some articles prevented us from making a more efficient use of the available information. In many study reports, data presentation was limited to the sample overall and did not present the results of the study according to customary subgroup stratification. Furthermore, there were many studies that had no obtainable data for pooling.

Appendix 5.—Statistical Methodology

The methods are taken from DerSimonian and Laird[10] and are based on work first presented by Cochran.[11] For combining information from k comparative clinical trials, each trial provides the number of patients in treatment and control groups, n_{ti} and n_{ci}, and the rate of some event in each of the two groups, r_{ti} and r_{ci}. In these analyses, the n_{ti} and n_{ci} are the minimum effective sample sizes obtained by dividing the number still at risk by the estimated success rates, namely, r_{ti} and r_{ci}, the life-table estimates obtained from the tables or curves in a publication. Let $RD_i = r_{ti} - r_{ci}$ denote the observed treatment effect for the i-th study, and assume that it is approximately $N(\theta_i, s_i^2)$, where θ_i is the true treatment effect for the i-th study, and s_i^2 is treated as known (and equal to $r_{ti}[1-r_{ti}]/n_{ti} + r_{ci}[1-r_{ci}]/n_{ci}$). The approximate normality assumption is reasonable when each observed effect is based on adequate sample sizes in the two groups. The studies being combined are regarded as a random sample from a population of studies. The population of treatment effects is modeled by assuming $E(\theta_i) = \mu$ and $Var(\theta_i) = \tau^2$, where μ estimates the overall treatment effect, τ^2 measures both the degree to which treatment effects vary across experiments as well as the degree to which individual studies give different assessments of treatment effects, and E indicates the expected value of the treatment effect.

The first step for pooling the treatment effects measured by the individual RD_i is to calculate the weighted mean:

$$\overline{RD} = \frac{\sum_{i=1}^{k} w_i \cdot RD_i}{\sum_{i=1}^{k} w_i}$$

where the weights are the inverse of the individual study variances: $w_i = s_i^{-2}$.

The next step involves calculating a measure of study heterogeneity, Q, as follows:

$$Q = \sum_{i=1}^{k} w_i (RD_i - \overline{RD})^2$$

If the study results are homogeneous, each separate study serves to estimate the same mean treatment effect (ie, all $\theta_i = \mu$), and the variance of the individual study means is zero (ie, $\tau^2 = 0$). If $\tau^2 = 0$, statistic Q is approximately χ^2 with $k-1$ *df*. Thus, by comparing the magnitude of Q with critical values from a $\chi^2(k-1)$ distribution, we have a statistical test for heterogeneity of the various studies. This statistical test has very low power and some adjustment for study heterogeneity seems reasonable whether or not a statistically significant value of Q is obtained.

To this end, the method of moments estimate for $Var(\theta_i) = \tau^2$ is calculated as follows:

$$\hat{\tau}^2 = \text{Max}\{0, (Q-[k-1])/(\Sigma w_i - \Sigma w_i^2/\Sigma w_i)\}$$

The weights for the final pooling step are then calculated as follows:

$$w_i^* = (s_i^2 + \hat{\tau}^2)^{-1}$$

The simple weighted method of moments estimates for μ and the SE of the estimate are as follows:

$$\hat{\mu} = \left(\sum_{i=1}^{k} w_i^* \cdot RD_i \right) \Big/ \sum_{i=1}^{k} w_i^*$$

and

$$SE(\hat{\mu}) = \left(\sum_{i=1}^{k} w_i^* \right)^{-1/2}$$

An approximate 95% confidence interval for the overall treatment effect rate difference is as follows:

$$\hat{\mu} \pm 1.96 \times SE(\hat{\mu})$$

To assess the impact of study quality on the results, we performed a sensitivity analysis to see whether the exclusion of particular low-quality[13] studies changed the results. Even when we excluded all long-term studies with an untreated control group whose internal validity score was lower than the median of 45%, the interpretation of the estimate was not changed. Likewise, the elimination of patients with negative nodes from the pooling of the first group of studies did not substantially alter the results in RFS or OS. Although only two of seven studies with quality below this threshold had data available for pooling (Table 1), biases in results favoring treatment did not appear to be substantially greater in the reported trials with lower internal validity score. Finally, we assessed the impact of the hormonal agent tamoxifen on pooling with the trial reporting the greatest effect (Ludwig group). We pooled data from this trial, comparing the CMFpT regimen with the pT treatment arm, rather than the untreated control arm. The pooled estimates obtained showed no substantial differences.

Some of the studies (at least in their published form) are still too recent to be fully evaluated. Hence, at the five-year point, only five of the eight studies available at three years could be pooled (Fig 2). Particularly with regard

Table 1.—Descriptive Characteristics of the 28 Randomized Control Trials of Adjuvant Chemotherapy for Breast Cancer (Listed in Each Group in Descending Order of Quality Score[s])

Study Name*	Sample Size	Regimen(s) Used†	Duration	Maximum Evaluable Time Point	Subgroups Reported‡	Significant Advantage Reported in: RFS§	Significant Advantage Reported in: OS§	Year First Reported
1. Long-term Chemotherapy With Untreated Control Group								
1. Ludwig	463‖	CMFpT pT‖	12 mo 12 mo	2 y	POST only; N1-3, N4+	POST; N1-3, N4+	NS	1984
2. NSABP (B-05)	349	L-Pam	24 mo	5 y	PRE, POST; N1-3, N4+	Overall; PRE	NS	1975
3. Milan	386	CMF	12 mo	8 y	PRE, POST; N1-3, N4+	Overall; PRE, N1-3	Overall	1976
4. Guy's II (CMF†)	327	CMF	12 mo	3 y	PRE, POST	Overall; PRE	NS	1984
5. ECOG¶	224	CMFPT CMFP	12 mo 12 mo	30 mo	POST only	POST at 1 y only	NS	1984
6. OSAKO#	240	LMF(+BCG)	6 mo (24 mo)	8 y	POST; N+, N−	Overall; N+; POST (includes N−)	NS	1978
7. Guy's I (L-Pam†)	370	L-Pam	22 mo	5 y	PRE, POST; N1-3, N4+	NS	NS	1983
8. MBCCG	252	CMFV	6 mo	3 y	PRE, POST	Overall; POST (at 12 mo)	NR	1977
9. Heidelberg group	100	LF	24 mo	Data too preliminary	None	NS	NR	1980
10. King's College	270	L-Pam	24 mo	Data not complete	None	NS	NS	1981
11. West Midlands Oncology Association	462	CMFAV (if N+)	6 mo	Data not complete	PRE, POST	Overall; PRE	NS	1979
	467	LMF (if N−)	6 mo	Data not complete	PRE, POST	NS	NR	1979
12. Glasgow	322	CMF ± XRT XRT	12 mo	1 y	PRE (OS§ only); N4+ (RFS§ only)	Overall; N4+	NS	1984
13. Ellis Fischel	167	Thiotepa	12 mo	5 y (OS§); 3 y (RFS§)	N−; N+ (RFS§ only)	N− only	NS	1969
14. Osaka group**	517	C-MmC ± XRT MmC ± XRT C ± XRT	6 mo 6 mo 6 mo	5 y	PRE, POST; N1-3, N4+	N1-3	NR	1980
2. Short-term Chemotherapy With Untreated Control Group								
1. Scandinavian Adjuvant Chemotherapy Study Group	1026	C (XRT to all patients)	6 d	10 y	PRE, POST; N1-3, N4+	Overall except where XRT delayed	Overall	1978
2. NSABP I (B-01)	820	Thiotepa or placebo	3 d	10 y	PRE, POST; N−, N1−3, N4+	PRE N4+ at 5 y only	PRE; PRE N4+ at 5 y only	1963
3. Newcastle	83	C	9 d	3 y	None	NS	NS	1971
3. Long-term Chemotherapy Comparing Different Regimens								
1. Western Cancer Study Group	62	CMF F	12 mo 12 mo	Data not complete	None	Only for first year	NS	1981
2. Southwest Oncology Group	361	CMFVP L-Pam	12 mo 24 mo	4 y	PRE, POST; N1-3, N4+	Overall; PRE, POST; N1-3, N4+	Overall; PRE; POST; N4+	1982
3. University of Alabama	171	CMF Oral L-Pam	12 mo 12 mo	Data not available	None	NS	Control group	1982
4. CALGB††	674	CMFVP CMF(±BCG)	22 mo 22 mo	3 y	N1-3; N4+; PRE, POST	N4+	NS	1982
5. Danish Breast Cancer Group‡‡	843‡‡ 155	CMF+XRT C+XRT XRT‡‡	12 mo 12 mo	1 y	PRE; PRE N0-3; PRE N4+	NS	NR	1981
6. Evanston Group	194	CFP (±BCG) L-Pam	12 mo 12 mo	Data too preliminary	None	Overall; POST at 1 y only	NR	1980

(Continued on p 1156.)

Table 1.—Descriptive Characteristics of the 28 Randomized Control Trials of Adjuvant Chemotherapy for Breast Cancer (Listed in Each Group in Descending Order of Quality Score[s]) (cont)

Study Name*	Sample Size	Regimen(s) Used†	Duration	Maximum Evaluable Time Point	Subgroups Reported‡	Significant Advantage Reported in: RFS§	Significant Advantage Reported in: QS§	Year First Reported
7. Mayo Clinic	293	CFP (±XRT) L-Pam	12 mo 12 mo	Data too preliminary	PRE, POST (RFS§ only)	PRE	PRE	1978
8. French Group	325	CAFV (if N+) CFV (if N−) CMF (N+ or N− control)	12 mo 12 mo 12 mo	5 y	PRE, POST	Overall; PRE N+; N+	NS	1984
9. Piedmont Oncology Association	158	CMF (±XRT) L-Pam (±XRT)	24 mo 24 mo	Data too preliminary	None	Overall	NR	1981
10. NASBP III (B-07)	741	L-PamF L-Pam	24 mo 24 mo	15 mo	PRE, POST; N1-3, N4+	POST	POST	1977
11. NASBP IV (B-08)	737	L-PamMF L-PamF	24 mo 24 mo	8 y	None	NS	NR	1980
4. Long-term Chemotherapy of One Regimen Given Over Different Duration§§								
1. Milan II	459	CMF CMF	12 mo 6 mo	5 y	PRE, POST; N1-3, N4+	PRE, POST; PRE N4+;' POST N1-3	NS	1979
2. SAKK*	400	LMF LMF	24 mo 6 mo	5 y	PRE, POST	NS	NS	1981
3. Southeastern Cancer Study Group	440	CMF CMF	12 mo 6 mo‖‖	Data not available	N1−3, N4+; PRE N1-3; POST N1-3; PRE N4+; POST N4+	PRE N1-3 at 42 mo	NR	1983

*NASBP indicates National Surgical Adjuvant Breast Project; ECOG, Eastern Cooperative Oncology Group; OSAKO, the East Switzerland Cooperative Oncology Group; MBCCG, Multicenter Breast Cancer Chemotherapy Group; CALGB, Cancer and Leukemia Study Group B; and SAKK, Swiss Association for Clinical Cancer Research.

†C indicates cyclophosphamide; M, methotrexate; F, fluorouracil; L-Pam, L-phenylalanine mustard; p, prednisone (low dose); P, prednisone (high dose); T, tamoxifen; L, chlorambucil; A, doxorubicin (Adriamycin); V, vincristine; MmC, mitomycin-C; BCG, Bacillus Calmette-Guérin; and XRT, radiation therapy.

‡PRE indicates premenopausal; POST, postmenopausal; N−, negative axillary lymph nodes; N+, positive axillary lymph nodes; N1-3, one to three axillary lymph nodes involved; and N4+, four or more axillary lymph nodes involved.

§RFS indicates relapse-free survival; OS, overall survival; NS, not statistically significant; and NR, not reported by authors.

‖Includes 164 patients randomized to pT.

¶The data for the two simultaneously randomized treatment arms were averaged arithmetically during pooling.

#The authors of the original study combined LMF with LMF+BCG-treated groups.

**Data not pooled since XRT was not assigned in randomized or systematic manner.

††Not eligible for data combination, since only patients with four or more nodes were eligible.

‡‡Not eligible for pooling because accrual to XRT control arm stopped in midtrial and data for two treated arms were too preliminary. Studied only premenopausal patients with chemotherapy regimens comparison.

§§This group of data was not extracted or combined.

‖‖Includes 71 patients with four or more involved nodes randomized to six months of CMF+XRT.

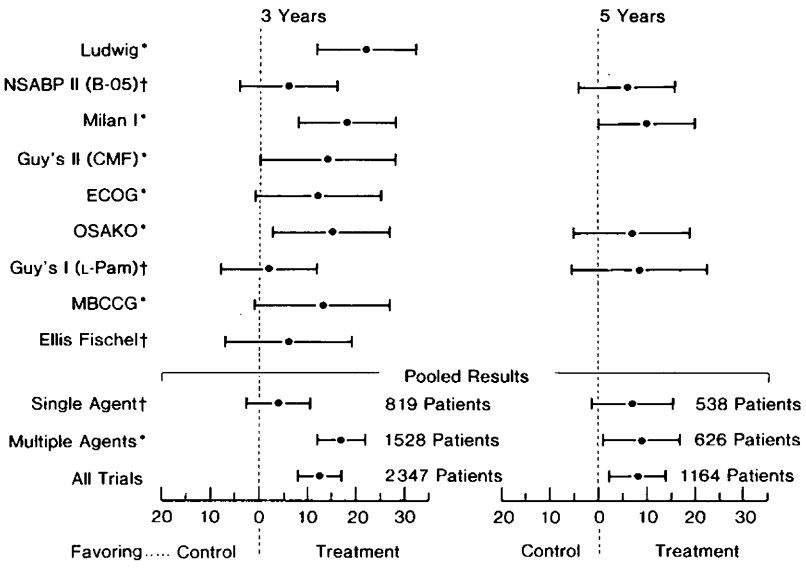

Fig 1.—Relapse-free survival expressed as observed rate differences at three and five years for long-term chemotherapy compared with an untreated control group. Daggers indicate single-agent regimens; asterisks, multiple-agent regimens; NSABP, the National Surgical Adjuvant Breast Project; CMF, cyclophosphamide-methotrexate-fluorouracil; L-Pam, L-phenylalanine mustard; ECOG, Eastern Cooperative Oncology Group; OSAKO, the East Switzerland Cooperative Oncology Group; and MBCCG, the Multicenter Breast Cancer Chemotherapy Group.

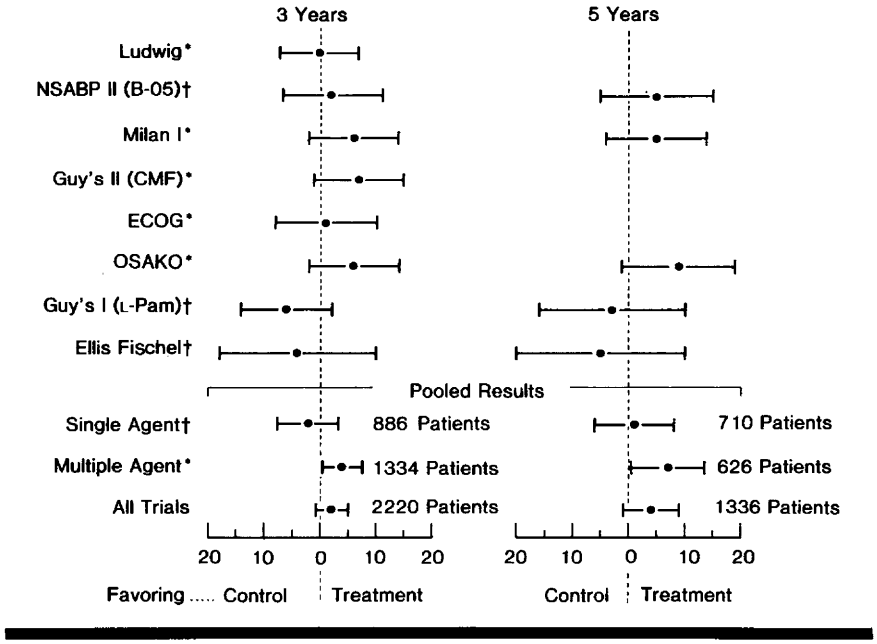

Fig 2.—Overall survival expressed as observed rate differences at three and five years for long-term chemotherapy compared with an untreated control group. Daggers indicate single-agent regimens; asterisks, multiple-agent regimens; NSABP, National Surgical Adjuvant Breast Project; CMF, cyclophosphamide-methotrexate-fluorouracil; L-Pam, L-phenylalanine mustard; ECOG, Eastern Cooperative Oncology Group; OSAKO, the East Switzerland Cooperative Oncology Group.

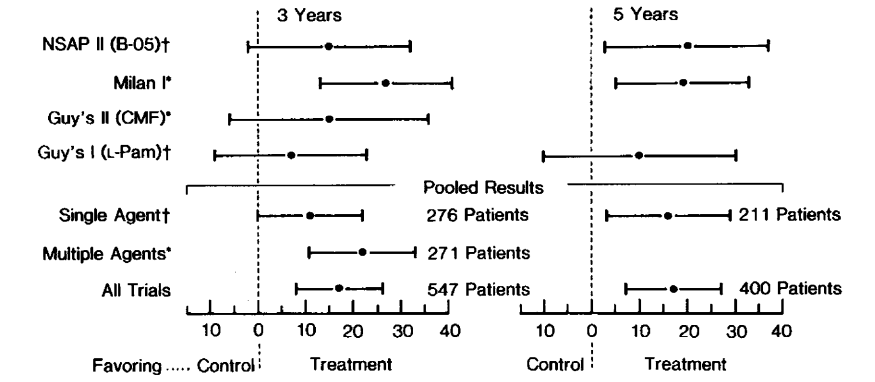

Fig 3.—Rate differences in relapse-free survival at three and five years among premenopausal patients. Daggers indicate single-agent regimens; asterisks, multiple-agent regimens; NSABP, National Surgical Adjuvant Breast Project; CMF, cyclophosphamide-methotrexate-fluorouracil; and L-Pam, L-phenylalanine mustard.

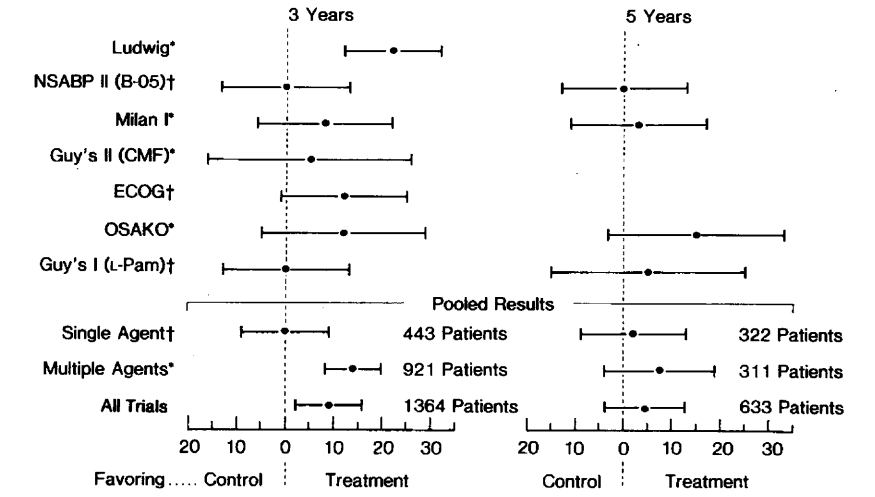

Fig 4.—Rate differences in relapse-free survival at three and five years among postmenopausal patients. Daggers indicate single-agent regimens; asterisks, multiple-agent regimens; NSABP, National Surgical Adjuvant Breast Project; CMF, cyclophosphamide-methotrexate-fluorouracil; L-Pam, L-phenylalanine mustard; ECOG, Eastern Cooperative Oncology Group; and OSAKO, the East Switzerland Cooperative Oncology Group.

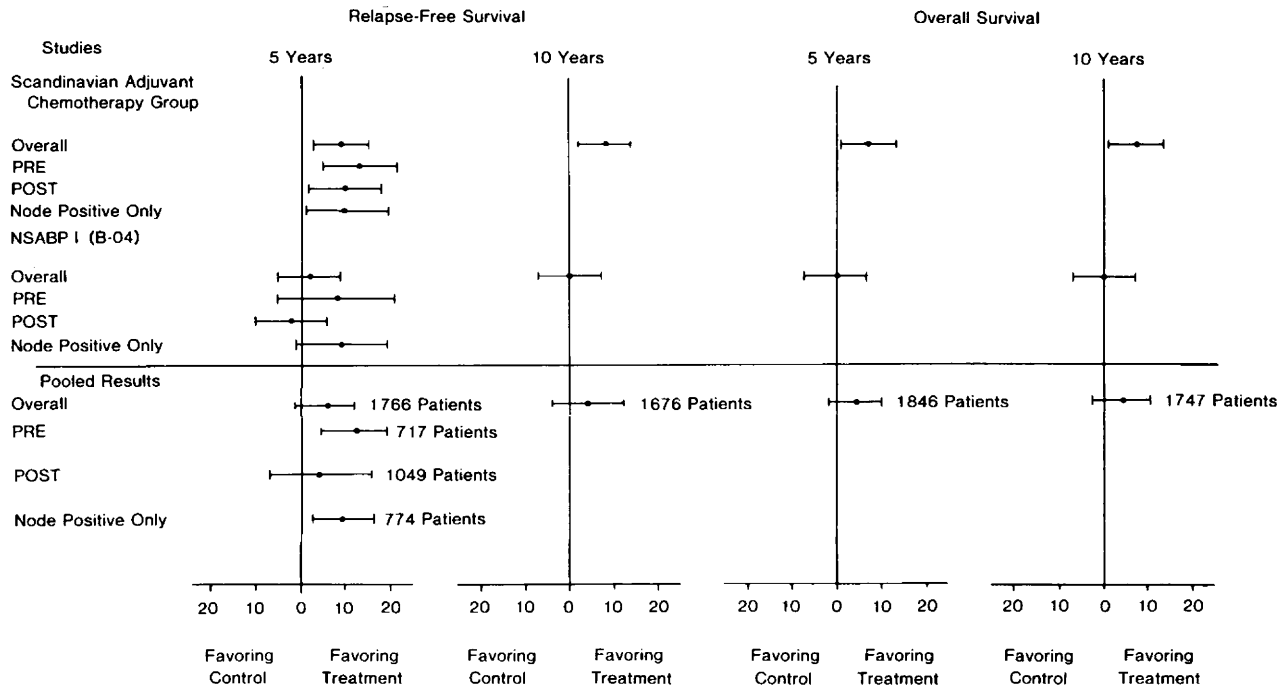

Fig 5.—Rate differences in relapse-free and overall survival in studies comparing short-term perioperative chemotherapy with no-treatment control groups. PRE indicates premenopausal; POST, postmenopausal; and NSABP, National Surgical Adjuvant Breast Project.

to OS, there may be a treatment benefit that is not currently apparent due to short follow-up or unavailability of the data.

A bias favoring the publication of positive trials may have influenced the results of this meta-analysis. An estimate of unpublished RCTs was obtained by reviewing the National Cancer Institute's directory of RCTs. There were six unpublished trials involving 1200 patients. Only one study with 180 patients was still in progress. The remaining five, including 1000 patients, showed no statistically significant treatment effect for adjuvant chemotherapy, except for the largest study of 450 patients, which showed improved RFS but not OS. In no study did the trend favor the control group.

The disparity in treatment effect between the RFS and OS raises a general question about the validity and clinical relevance of RFS as an end point. In another study,[13] we demonstrated that only 28% of the RCTs reported a specific and well-defined schedule for the follow-up of study subjects. This lack of standardization among treatment protocols might account for some of the variation in RFS between studies. Furthermore, the lack of blinding of observers makes this end point susceptible to bias even when the follow-up is standardized in a trial. Whether or not delay of RFS is consistently followed by prolonged survival requires the follow-up of many more randomized patients.

Critics of data pooling might raise objections to the combination of data as presented herein. However, any clinician who considers more than one RCT's results to be valid and noteworthy will be combining data, albeit subconsciously and implicitly, when deciding about the proper treatment for a patient. Moreover, reviewers generally consider only those studies that appear to be worthwhile and discount those RCTs (and their conclusions) that are not in agreement with a favored study or group of studies. This usual method of policymaking amounts to implicitly assigning weights to disparate studies according to a nonsystematic and biased evaluation scheme and combining them in a nonquantitative manner. The resulting conclusions are individualized and nonrepeatable, with reasons for exclusion or downweighting of data rarely explicitly stated.

In this analysis, all data available from each study have been extracted into a standardized format and given weight according to sample size and heterogeneity. The result represents a systematic, quantitative pooling of the data available in published reports on adjuvant chemotherapy for primary breast cancer. Similarly, the availability of more recent results may easily be incorporated into this method to continually update our findings.

The choice of methods used raises two statistical issues. Pooling methods could have been selected based on the relative risk or the ratio of odds of treatment failure. The rate difference was chosen because experience with breast cancer data indicates the risk of failure of treatment is not constant and is unlikely to be proportional over time. Thus, a study with short-term follow-up may indicate a large treatment effect in terms of the odds of failure, but the same study may demonstrate no effect when more mature. Three-, five-, and ten-year time points were chosen because the three-year point was most commonly reported and the five- and ten-year points are considered routine in end-result reporting.

The second statistical consideration in pooling involves the choice of mathematical model used to relate the observed results to the theoretical underlying "true" treatment effect. The selection of the DerSimonian and Laird random effects model[10,11] is based

on the concept that the individual studies cannot be expected to provide observed results that are realizations of the same distribution. Each study has individual characteristics of patient eligibility, referral patterns, experimental environment, treatment applied, follow-up policy, measurement of treatment effect, and reporting of criteria for excluding and rejecting patients from the study. All of these characteristics affect the types and degrees of bias influencing the observed treatment effect. The random-effects model allows for the fact that studies each have their own underlying treatment difference (denoted θ_i) that are themselves representative of a superpopulation having an overall mean treatment effect denoted by μ. If an overall effect exists for the treatment, μ will be greater than 0, and the pooled estimate of μ observed can be used to test the hypothesis that $\mu=0$. The value of the method is that not only does it test for study homogeneity with the Q statistic, but it also incorporates the degree of study heterogeneity into the estimate of the variability of the pooled estimate of treatment effect. In contrast to other methods for pooling data,[17-20] the DerSimonian-Laird method enlarges the confidence interval for the overall estimate if study heterogeneity is large.

More critical than the actual rates determined by this study is the fact that the results of the pooled estimates derived are universally accessible to any reader and can be confirmed by repeating the calculations performed. In fact they can be recalculated according to any assessment of quality or other scale for judging data eligibility. It is hoped that this method will be viewed as a preliminary approach to solving a therapeutic controversy in medicine where the RCTs that have been reported are inconclusive when considered individually. Furthermore, the disheartening lack of derivable data is presented in the hope that it will stimulate an improvement in the publication standards for medical reporting.

This study was supported in part by grant LM31116 from the National Library of Medicine to Mount Sinai School of Medicine.

Dr Himel was a W. K. Kellogg Foundation Fellow in Health Policy and Management on leave of absence from the Boston University Medical Center. Dr Liberati was a fellow of the Italy-US Association for Cancer Research, on leave of absence from the Mario Negri Institute of Pharmacologic Research, Milan, Italy. Dr Gelber is Associate Professor of Biostatistics, Harvard School of Public Health and Dana-Farber Cancer Institute, Boston. Dr Chalmers was a Visiting Professor at the Harvard School of Public Health.

We would like to thank Elgie Ginsbergh and Peg Hewitt, MLS, for their technical assistance in performing library searches; Karen Uliss Abbett and Tabitha Payne for their assistance in the preparation of this manuscript; and Chung-Cheng Hsieh, SD, for programming the algorithm used to perform the arithmetic computations.

References

1. NIH Consensus Development Statement: Adjuvant chemotherapy of breast cancer. *N Engl J Med* 1980;303:831-832.
2. Fisher B, Redmond CK, in collaboration with other NSABP investigators: Topics in oncology: Advances in adjuvant chemotherapy of breast cancer, in Fairbanks VF (ed): *Current Hematology.* New York, John Wiley & Sons Inc, 1983, vol 2, pp 415-446.
3. Bonadonna G, Valagussa P, Veronesi U: Current views on the primary treatment of breast cancer, in Burchenthal JH, Oetlgen HS (eds): *Cancer: Achievements, Challenges and Prospects for the 1980's.* New York, Grune & Stratton Inc, 1981, pp 337-356.
4. Goldhirsch A, Gelber RD, Davis BW: Adjuvant chemotherapy trials in breast cancer: An appraisal and lessons for patient care outside the trials, in Forbes JF (ed): *Clinical Surgery International—10: Breast Disease.* Edinburgh, Churchill Livingston, 1985, pp 123-138.
5. Cooper RG: Adjuvant chemotherapy and the practicing oncologist. *Surg Clin North Am* 1984; 64:1173-1179.
6. Carter SK: Adjuvant chemotherapy for breast cancer: Implications of clinical trials. *Postgrad Med* 1985;77:75-83.
7. Blamey RW: Adjuvant chemotherapy in breast cancer: The case against, in Delaney J, Varco R (eds): *Controversies in Surgery II.* Philadelphia, WB Saunders Co, 1983, pp 94-103.
8. Rubens RD, Knight RK, Fentiman IS, et al: Controlled trial of adjuvant chemotherapy with melphalan for breast cancer. *Lancet* 1983;1:839-843.
9. Liberati A, Patterson WB, Biener L, et al: The impact of early breast cancer randomized control trials on the practice of medicine. Abstract of the Sixth Annual Meeting of Society for Clinical Trials, New Orleans, May 13-15, 1985.
10. DerSimonian R, Laird N: Meta-analysis and clinical trials. *Controlled Clin Trials,* 1986, vol 7, No. 3.
11. Cochran WG: The combination of estimates from different experiments. *Biometrics* 1954; 10:101-129.
12. Goldman L, Feinstein AR: Anticoagulants and myocardial infarction: The problems of pooling, drowning, and floating. *Ann Intern Med* 1979; 90:92-94.
13. Liberati A, Himel HN, Chalmers TC: A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol* 1986;4:942-951.
14. Chalmers TC, Smith H Jr: A method for assessing the quality of a randomized control trial. *Controlled Clin Trials* 1981;2:31-49.
15. Ludwig Breast Cancer Study Group: Toxic effects of early adjuvant chemotherapy for breast cancer. *Lancet* 1983;2:542-544.
16. Cancer Studies Unit Conference: Review of mortality results in randomised trials in early breast cancer. *Lancet* 1984;2:1205.
17. Rosenthal R: Combining results of independent studies. *Psych Bull* 1978;85:185-193.
18. Canner P: Aspirin and coronary disease: Comparison of six clinical trials. *Isr J Med Sci* 1983;19:413-423.
19. Stampfer MJ, Goldhaber SZ, Yusuf S, et al: Effect of intravenous streptokinase on acute myocardial infarction: Results from pooled randomized trials. *N Engl J Med* 1982;307:1180-1182.
20. Yusuf S, Peto R, Lewis J, et al: Beta-blockade during and after myocardial infarction: An overview of the randomized trials. *Prog Cardiovasc Dis* 1985;27:335-371.