Effectiveness of an Adjuvant Chemotherapy Regimen for Early-Stage Breast Cancer
A Systematic Review and Network Meta-analysis

Takeo Fujii, MD, MPH; Fanny Le Du, MD; Lianchun Xiao, MS; Takahiro Kogawa, MD, PhD; Carlos H. Barcenas, MD, MS; Ricardo H. Alvarez, MD, MS; Vicente Valero, MD; Yu Shen, PhD; Naoto T. Ueno, MD, PhD

**IMPORTANCE** Different adjuvant chemotherapy regimens are available for early-stage breast cancer. Because conventional meta-analysis does not allow comparing all regimens, we performed a network meta-analysis to identify the most effective adjuvant chemotherapy regimen.

**OBJECTIVE** To find the most effective adjuvant therapy regimen for early-stage breast cancer.

**DATA SOURCES** We searched MEDLINE, Embase, and the Cochrane Library for articles published before June 2015; the American Society of Clinical Oncology annual meeting abstracts from January 1983 through December 2014; and the American Association for Cancer Research annual meeting abstracts from January 1916 through December 2014. Additionally, we manually searched bibliographies for related references.

**STUDY SELECTION** We included randomized clinical trials of adjuvant treatments for early-stage breast cancer that compared 2 or more of the following: no adjuvant chemotherapy; sequential anthracycline-cyclophosphamide and taxane (AC-T); concurrent anthracycline-cyclophosphamide and taxane (ACT); anthracycline-cyclophosphamide without taxane (AC); docetaxel and cyclophosphamide (TC); cyclophosphamide, methotrexate, and fluorouracil (CMF); and platinum-containing regimens.

**DATA EXTRACTION AND SYNTHESIS** We followed the PRISMA guidelines. Two investigators independently selected the articles and extracted information. Disagreements were resolved by discussion with another author. Quality was assessed by Cochrane risk-of-bias method. Data were pooled using random-effects models.

**MAIN OUTCOMES AND MEASURES** We used network meta-analysis to test the most effective adjuvant therapy regimen in terms of overall survival (OS) by comparing regimens listed in the National Comprehensive Cancer Network guidelines and platinum-containing regimens.

**RESULTS** We identified 24 trials. The TC and platinum-containing regimens had OS benefit similar to that of sequential AC-T (TC hazard ratio [HR], 0.93; 95% CI, 0.62-1.40; and platinum HR, 0.93; 95% CI, 0.66-1.31). Patients treated with CMF or AC had significantly worse OS than those treated with sequential AC-T (CMF HR, 1.56; 95% CI, 1.32-1.85; and AC HR, 1.22; 95% CI, 1.10-1.37). Platinum-containing regimens tended to be more toxic than sequential AC-T. The toxicity of TC was similar to or less than that of sequential AC-T. Meta-regression analysis showed that hormone receptor status did not impact the HRs for OS for any regimen.

**CONCLUSIONS AND RELEVANCE** Sequential AC-T is likely to be the most effective adjuvant therapy regimen for early-stage breast cancer regardless of hormone receptor status.
Breast cancer is one of the most common cancers in women worldwide. In the United States, breast cancer is the most common cancer in women and the second most common cause of cancer death in women. It is well established that adjuvant chemotherapy plays an important role in reducing the risk of recurrence and improving the survival of patients with breast cancer.

The National Comprehensive Cancer Network (NCCN) guidelines for the treatment of breast cancer describe numerous recommended adjuvant chemotherapy regimens, including sequential anthracycline-cyclophosphamide and taxane (AC-T); concurrent anthracycline-cyclophosphamide and taxane (ACT); anthracycline-cyclophosphamide without taxane (AC); cyclophosphamide, methotrexate, and fluorouracil (CMF); and docetaxel and cyclophosphamide (TC). Among them, sequential AC-T is the most commonly accepted standard regimen. However, 2 types of regimens without anthracyclines—TC and platinum-containing regimens—may have efficacy similar to or greater than that of sequential AC-T. A previous study showed that TC was superior to doxorubicin and cyclophosphamide with regard to disease-free survival and overall survival (OS) and was less toxic. Platinum-containing regimens have also demonstrated high efficacy against breast cancer with tolerable adverse events (AEs).

However, those regimens have never been compared with sequential AC-T because of limitations of conventional meta-analysis and the lack of direct comparison trials. Thus, the most effective adjuvant chemotherapy regimen for early-stage breast cancer is still not known.

Network meta-analysis is the best way to identify that regimen. By using indirect comparisons, network meta-analysis allows comparisons of treatments for which there have been no head-to-head comparison. For example, we can compare outcome data about treatment A and treatment B derived from multiple studies that included either treatment A or treatment B and a common treatment C. Network meta-analysis allowed us to compare TC with sequential AC-T and compare platinum-containing regimens with sequential AC-T.

Using network meta-analysis in this study, our primary objective was to find the most effective adjuvant therapy regimen for early-stage breast cancer in terms of OS by comparing regimens listed in the NCCN guidelines and platinum-containing regimens. Our secondary objectives were to determine event-free survival and AEs.

Methods

Literature Search and Selection Criteria

We performed a systematic review of the literature according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. We searched MEDLINE, Embase, and the Cochrane Library for articles published before June 1, 2015; the American Society of Clinical Oncology annual meeting abstracts from January 1983 through December 2014; and the American Association for Cancer Research annual meeting abstracts from January 1916 through December 2014. The search terms used were “breast neoplasms,” “chemotherapy, adjuvant,” “randomized controlled trial,” “controlled clinical trial,” “random allocation,” “double-blind method,” “single-blind method,” “survival analysis,” “treatment outcome,” “anthracyclines,” “taxoids,” “cyclophosphamide,” “fluorouracil,” “methotrexate,” “cisplatin,” and “carboplatin.” Additionally, we manually searched bibliographies and added related references. We included only randomized clinical trials of adjuvant treatments for early-stage breast cancer and added related references. We included only randomized clinical trials of adjuvant treatments for early-stage breast cancer that compared 2 or more of the following options: no adjuvant chemotherapy, sequential AC-T, concurrent ACT, AC, TC, CMF, and platinum-containing regimens. Early-stage breast cancer was defined as pathological stage I to III. We excluded studies if they did not provide enough data to obtain hazard ratios (HRs) for survival. Trials for which full-text reports were not available were also excluded.

Data Collection and Quality Assessment

A well-trained librarian performed a comprehensive literature search. Two investigators (T.F., F.L.) independently reviewed and selected the articles and extracted information. Any disagreements were resolved by discussion with another author (T.K.). We collected the number of randomized patients, the number of analyzed patients, treatment regimens, median follow-up time, median age, age range, HRs for OS and event-free survival (EFS), and the number of unacceptable AEs, which were defined as AEs grade 3 or greater that can require dose reduction or schedule delay. To identify hematologic unacceptable AEs, we counted the number of cases of leukocytopenia when the differentials were not reported, neutropenia, anemia, thrombocytopenia, febrile neutropenia, and secondary hematologic malignancies such as leukemia and myelodysplastic syndrome. To identify nonhematologic unacceptable AEs, we counted the number of cases of diarrhea, cardiotoxicity, hypersensitivity reactions, neurotoxicity, thromboembolism, nausea, vomiting, anorexia, asthenia, mucositis, any death, and elevations of transaminase, total bilirubin, and creatinine.

In 2 trials, only the total number of randomized patients was provided, and we assumed that the patients were equally distributed into each treatment group because these 2 trials used 1-to-1 randomization. The effect on the analysis results...
must be small.14,15 For AEs, we calculated ORs by approximating the number of patients experiencing unacceptable AEs by the number of AEs (ie, assuming 1 adverse event per patient). If only percentages were reported, we calculated the number of AEs by multiplying percentages by the number of patients. When the percentages were reported as less than 1% and the actual percentages were not available, we used 0.5% and rounded the result up to the nearest whole number.8,16 We excluded 7 trials from toxicity analysis because of insufficient information.14,17-22 In 1 trial, only grade 4 hematologic AEs were reported.16

Risk of bias of individual studies was assessed by using the Cochrane risk-of-bias method.23 Studies are assessed on the basis of sequence generation, allocation concealment, blinding, incomplete outcome data, selective reporting, and other sources of bias. Blinding was not performed in all the trials, but assessment of survival was not likely to be influenced by lack of blinding.

Institutional review board review and informed consent were waived because this analysis was based on information from published articles.

**Statistical Analyses**

The primary outcome was OS, and the secondary outcomes were EFS and the number of unacceptable AEs. To match the definitions used in the original studies, we defined OS as the time from date of randomization, first dose of adjuvant chemotherapy, or surgery until death from any cause. We defined EFS as the time from date of randomization, first dose, or surgery until recurrence, contralateral breast cancer, second primary cancer, or death from any cause, whichever occurred first.

Hazard ratios and the corresponding 95% CIs were extracted from the study reports. If HRs and 95% CIs were not reported, we estimated the HRs either using the reported median OS and EFS times and P values from log-rank tests or by using OS rates and the number of events in each arm at a given time (eg, 5 years). Specifically, the natural logarithm (ln) of HR (ie, ln (HR)) and the standard error (SE) of the ln multiplied by the HR (ie, SE (ln (HR))) were computed using the summary statistics.24,25 Then the pertinent graph-theoretical method was applied to perform network meta-analysis using the transformed HRs and the corresponding SEs from different studies.12 Random-effects models were used to account for heterogeneity between studies.26 To assess the heterogeneity, we calculated Cochran Q statistics and inconsistency statistics (I²).

A prior report showed that I² index values of less than 25% indicate low heterogeneity; 25% to 50%, medium heterogeneity; and greater than 50%, high heterogeneity.27 Descriptively, a net heat plot derived from the Q statistic of the network was provided for each of the 2 outcomes. The net heat plot is a matrix plot to assess any inconsistency, which represents the influence of the detachment of each treatment comparison on the rest of the comparisons in the network. Inconsistency in this context means disagreement between direct and indirect therapy comparisons in addition to difference among the same treatment arms by studies.28 More intense color (eg, red) indicates an inconsistent treatment comparison relative to the network, and less intense color (eg, yellow) indicates a consistent treatment comparison relative to the network. A generalized linear model with Bayesian framework was fit to compare treatment effects on OS or on EFS adjusting for study level hormone receptor status.29

We used R, version 3.0.2, and STATA, version 13 (STATA Corp), for analyses.

**Results**

The details of our literature search are shown in eFigure 1 in the Supplement. Of 4324 potentially relevant articles, 31 articles (24 trials) were included in this study.7,8,14-22,30-49 The results of the quality assessment of the trials according to the Cochrane risk-of-bias tool are shown in eTable 1 in the Supplement. Twenty-one trials included only women and in 3 trials there was not a clear description of criteria about sex.10,21,42 We believe that this does not affect the results because the majority of patients with breast cancer are women.

The 24 trials are summarized in eTable 2 in the Supplement. In the network plot in Figure 1, node size is proportional to the total number of patients in the treatment group. Line width is proportional to the number of trials comparing the treatment groups connected by the line, which are represented by the numbers next to each line. A total of 24 trials were analyzed. AC indicates anthracycline-cyclophosphamide without taxane; AC-T, sequential anthracycline-cyclophosphamide and taxane; ACT, concurrent anthracycline-cyclophosphamide and taxane; CMF, cyclophosphamide, methotrexate, and fluorouracil; TC, docetaxel and cyclophosphamide.
baseline patient and disease characteristics are summarized in eTable 3 in the Supplement. Details of the treatment regimens are presented in eTable 4 in the Supplement. In total, our analysis included 28,853 patients: 8109 treated with sequential AC-T; 6039, concurrent ACT; 10,781, AC; 506, TC; 1935, CMF; 255, platinum-containing regimens; and 1228, no adjuvant chemotherapy.

In all analyses, we used sequential AC-T as the reference because sequential AC-T is the most commonly accepted standard regimen for adjuvant chemotherapy for early-stage breast cancer. In OS analysis, we included 21 trials. For 19 trials, HRs were reported in the articles. For the other 2 trials, we estimated the HRs on the basis of information provided.21 The HRs are shown in Figure 2B and eTable 7 in the Supplement. Sequential AC-T yielded longer OS than AC, CMF, or no adjuvant chemotherapy. No statistically significant difference was detected between sequential AC-T and concurrent ACT, platinum-containing regimens, or TC. Cyclophosphamide, methotrexate, and fluorouracil; TC, docetaxel and cyclophosphamide.

In EFS analysis, we included 22 trials. In 1 trial, HR was not reported, and we estimated it on the basis of information provided.21 The HRs are shown in Figure 2B and eTable 7 in the Supplement. Sequential AC-T was associated with better EFS than AC, CMF, and no adjuvant chemotherapy. No statistically significant difference was detected between sequential AC-T and concurrent ACT, platinum-containing regimens, or TC. Sensitivity analysis was performed with only the 20 studies for which HRs were reported in the original articles (eTable 8 in the Supplement). The results were similar to those for the analysis of all 22 studies. Meta-regression analysis on EFS adjusted for the percentage of hormone receptor-positive patients in each study showed HRs similar to those without adjustment for this study-level factor (Figure 3B).

Based on the rank analysis order, the worst adjuvant regimen based on OS to the best is: no adjuvant chemotherapy, CMF, AC, concurrent ACT, sequential AC-T, platinum-containing regimens, and TC. The probability to be the best was 43% for TC and 39% for platinum-containing regimens (eTable 9 in the Supplement). However, those 2 regimens did not show significant difference in OS and EFS (Figure 2).

The Q statistic was 18.17 (P = .25) for the OS analysis and 15.08 (P = .52) for the EFS analysis. The I² index was 17.5% for the OS analysis and 0% for the EFS analysis. The Q statistic informs us about presence or absence of heterogeneity. The I² index is to quantify the extent of heterogeneity. When the P value of Q statistics is greater than or equal to 0.05, there is not enough evidence to reject the homogeneity assumption. When the I² index is less than 25%, it means low heterogeneity. Both Q statistics and I² confirmed that heterogeneity was low. To provide a better understanding of the inconsistency
of treatment comparison visually, the net heat plots for survival outcome were applied. Consistent with the Q statistics, the net heat plots showed mild colors for all the comparisons, which indicated that no significantly inconsistent treatment comparisons influenced the network meta-analysis (eFigure 2 in the Supplement).

In the analysis of unacceptable AEs, we excluded 7 trials because we could not attain the necessary information.14,17-21 For the 17 trials, network meta-analysis was performed separately for overall, hematologic, and nonhematologic unacceptable AEs.

The estimated odds ratios (ORs) for overall unacceptable AEs for pairwise comparisons between different treatment regimens are shown in Figure 4A and eTable 10A in the Supplement. Platinum-containing regimens tended to have more unacceptable AEs than sequential AC-T (OR, 3.55; 95% CI, 0.8-15.71) (Figure 4A). Cyclophosphamide, methotrexate, and fluorouracil was associated with fewer overall unacceptable AEs than sequential AC-T, concurrent ACT, or platinum-containing regimens.

The estimated ORs for hematologic unacceptable AEs are shown in Figure 4B and eTable 10B in the Supplement. Concurrent ACT tended to have more hematologic unacceptable AEs than other regimens. Cyclophosphamide, methotrexate, and fluorouracil was less toxic than the other regimens.

The estimated ORs for nonhematologic unacceptable AEs are shown in Figure 4C and eTable 10C in the Supplement. Platinum-containing regimens had more unacceptable AEs than sequential AC-T (OR, 5.17; 95% CI, 1.36-19.67) (Figure 4C), and CMF showed a tendency toward being associated with fewer AEs than the other regimens.

For all of the overall, hematologic, and nonhematologic unacceptable AEs, there were no significant differences between TC and sequential AC-T (overall AEs OR, 0.56; 95% CI, 0.13-2.35; hematologic AEs OR, 1.08; 95% CI, 0.37-3.17; nonhematologic AEs OR, 0.32; 95% CI, 0.09-1.12) (eTable 10 in the Supplement).

The strength of evidence of our current study was evaluated based on the US Agency for Healthcare and Research Quality’s approach (eTable II in the Supplement). Directness of some comparisons was evaluated as indirect. However, this indirect comparison by using network meta-analysis is one of our strong points and can provide HRs of treatments for which there has not been head-to-head comparison.

Discussion

We draw several important conclusions from this study. First, sequential AC-T should still be the first choice for chemotherapy in the general population of patients with early-stage breast cancer on the basis of OS and risk of unacceptable AEs. Our meta-regression analysis to consider the potential effect of hormone receptor status on OS showed that findings were similar after adjustment for hormone receptor status. Second, TC was similar to sequential AC-T in terms of treatment effect and unacceptable AEs and might be considered as a first choice treatment for patients with high risk of cardiotoxic effects. Third, although platinum-containing regimens have recently been considered for breast cancer treatment, they are not superior to sequential AC-T in terms of OS and tend to be more toxic.

To our knowledge, this is the first network meta-analysis to provide estimates of HRs for OS and EFS for pairwise comparisons of potential adjuvant chemotherapy regimens for patients with early-stage breast cancer. We can rank efficacy by using these pooled HRs.

In our study, we included platinum-containing regimens because some prior trials showed their efficacy, but the efficacy of platinum-containing regimens in the general population of patients with early-stage breast cancer is still unknown.50,51 Historically, combination therapy with an anthracycline and a taxane resulted in better clinical outcomes than CMF or anthracycline regimens without taxane, but no study had compared all of these regimens simultaneously because of the limitations of conventional statistical methods.

Regarding the quality of selected trials, the kinds of AEs reported in each trial are different and selective reporting must exist. However, we only included unacceptable AEs, which we believe should be reported to minimize the effect of positive selective reporting. In 7 of 24 trials, sequence generation was unclear, and in another 7 of 24 trials, allocation concealment was unclear. However, the patient characteristics in each trial were equally distributed among each regimen.
HRs for OS on the basis of the reported survival rates and the in 3 studies, HRs were not provided, and we had to estimate short compared with the long expected OS duration. Second, randomization, surgery, and first dose of chemotherapy should be very adjuvant chemotherapy, intervals between dates of randomization, surgery, and first dose of chemotherapy should be very short compared with the long expected OS duration. Second, in 3 studies, HRs were not provided, and we had to estimate HRs for OS on the basis of the reported survival rates and the log-rank test P values or the number of events. However, the sensitivity analysis excluding those trials produced results similar to those of the overall analysis.

Our trial has limitations related to data collection. We grouped several different sequential anthracycline and taxane-based regimens, including AC followed by weekly paclitaxel, AC followed by docetaxel, AC followed by paclitaxel every 3 weeks, and dose-dense AC followed by paclitaxel. Also regimens with different doses of each drug were grouped together. Another limitation is that clinical trials do not necessarily reflect the real-world toxic effects of these regimens. For example, TC was later reported to be associated with a very high rate of neutropenic fever not seen in the original clinical trials. Because all of the regimens are combination chemotherapies, how much each drug causes AEs cannot be determined. Some trials were conducted many years ago, and subdata (eg, HR information) were not collected. However, since this study included only randomized trials, the distributions of patient and disease backgrounds should be similar. One of the biggest findings supporting patient and disease background similarity is that the HRs for OS and EFS in the meta-regression analysis, adjusting for hormone receptor status at the study level, were not different from those without the adjustment. This meta-regression result can potentially be subject to the ecological fallacy because each study did not report HRs comparing patients with and without hormone receptor positivity. Human epidermal growth factor receptor 2 (HER2) status was available in only 4 trials, and we could not adjust for it because of the limited information. However, its effects must be small given that we included only randomized trials. Median patient age in this study is relatively young; only 1 study had an average patient age older than 50 years. This may affect the probability of AEs (ie, cardiotoxic effects). However, because we only included randomized clinical trials, the effect of the incidence probability difference must be small.

We excluded trials targeting HER2-positive patients to compare the effect of cytotoxic agents. This might be one of the limitations because anti-HER2 target therapy in patients who are HER2 positive may have a significant effect on survival outcomes. Additionally, the median follow-up duration might have some effect on survival analysis. In one trial, median follow-up was not reported. Median follow-up of 4 trials was less than 60 months (2 trials, 58 months; 1 trial, 59 months; and 1 trial, 55 months). Also, the follow-up schedule would vary depending on the trials, and the detection of an event may be biased depending on the frequency and the methods of follow-up. Because this is a network meta-analysis using data drawn from published articles and because the sample size of life-threatening events and long-term AEs is limited, we cannot draw a statistically meaningful conclusion when we analyze those AEs independently.

Recently, platinum agents have been suggested to be effective. However, we detected no significant difference in terms of OS or EFS between sequential AC-T and platinum-containing regimens. Platinum-containing regimens were associated with a higher risk of toxic effects than sequential AC-T. Although the number of trials included in the analysis was limited and the treatment dose might be high compared with the dose of carboplatin used in a docetaxel, carboplatin, and trastuzumab regimen, which can lead to more toxic effects than platinum-containing regimens, we do not recommend platinum-containing regimens unless we can select a specific target population that is more sensitive to them. Currently, studies are investigating target markers that may be predictive of good responses to platinum agents, especially in patients with triple-negative breast cancer and/or breast cancer with BRCA mutations. Furthermore, we found that TC was similar to sequential AC-T in terms of OS and unacceptable AEs. The TC regimen remains useful for patients with a high risk of cardiotoxicity.

**Conclusions**

We recommend sequential AC-T regimens for the general population of patients with early-stage breast cancer. In terms of treatment effect and unacceptable AEs, TC might also be acceptable, especially for patients with a high risk of cardiotoxic effects. Platinum-containing regimens should not be recommended to the general population of patients with breast cancer. Further clinical studies may be required to investigate target markers that will predict patients with a higher likelihood of response to platinum-containing regimens. Sequential AC-T is likely to be the most effective adjuvant treatment for early-stage breast cancer regardless of hormone receptor status.
Role of the Funder/Sponsor: The Morgan Welch Inflammatory Breast Cancer Research Program, the State of Texas Rare and Aggressive Breast Cancer Research Program, and the National Cancer Institute (grant CA079466), which supports the Biostatistics Shared Resource, and the National Cancer Institute (grant CA079466).

Additional Contributions: Stephanie P. Deming, BA, of the Department of Scientific Publications at The University of Texas MD Anderson Cancer Center provided scientific editing services. Yinmin Geng, MS, MLS(is), clinical librarian of the Research Medical Library at The University of Texas MD Anderson Cancer Center, conducted the literature search. Ms Deming and Ms Geng were both compensated for their contributions.

REFERENCES


