OXFORD

ARTICLE

Timing of Adjuvant Surgical Oophorectomy in the Menstrual Cycle and Disease-Free and Overall Survival in Premenopausal Women With Operable Breast Cancer

Richard R. Love, Adriano V. Laudico, Nguyen Van Dinh, D. Craig Allred, Gemma B. Uy, Le Hong Quang, Jonathan Disraeli S. Salvador, Stephen Sixto S. Siguan, Maria Rica Mirasol-Lumague, Nguyen Dinh Tung, Noureddine Benjaafar, Narciso S. Navarro Jr., Tran Tu Quy, Arturo S. De La Peña, Rodney B. Dofitas, Orlino C. Bisquera Jr., Nguyen Dieu Linh, Ta Van To, Gregory S. Young, Erinn M. Hade, David Jarjoura

Affiliations of authors: International Breast Cancer Research Foundation, Madison, WI (RRL); University of the Philippines Manila, Philippine General Hospital, Manila, Philippines (AL, GU, ADP, RD, OB); Hospital K, Hanoi, Vietnam (NVD, LHQ, NDL, TVT); Clarient Pathology Services, Aliso Viejo, CA (DCA); East Avenue Medical Center, Manila, Philippines (JS); Vicente Sotto Hospital, Cebu, Philippines (SSS); Rizal Medical Center, Manila, Philippines (RML, MHA); Hue Central Hospital, Hue, Vietnam (NDT); National Institute of Oncology, Rabat, Morocco (NB); Santo Tomas University Hospital, Manila, Philippines (NN); Danang General Hospital, Danang, Vietnam (TTQ); The Ohio State University Center for Biostatistics, Columbus, OH (GSY, EMH, DJ (ret.).

Correspondence to: Richard R. Love, MD, 2708 Columbia Road, Madison, WI 53705 (e-mail: richardibcrf@gmail.com).

Abstract

Background: For women with hormone receptor–positive, operable breast cancer, surgical oophorectomy plus tamoxifen is an effective adjuvant therapy. We conducted a phase III randomized clinical trial to test the hypothesis that oophorectomy surgery performed during the luteal phase of the menstrual cycle was associated with better outcomes.

Methods: Seven hundred forty premenopausal women entered a clinical trial in which those women estimated not to be in the luteal phase of their menstrual cycle for the next one to six days (n = 509) were randomly assigned to receive treatment with surgical oophorectomy either delayed to be during a five-day window in the history-estimated midluteal phase of the menstrual cycles, or in the next one to six days. Women who were estimated to be in the luteal phase of the menstrual cycle for the next one to six days (n = 231) were excluded from random assignment and received immediate surgical treatments. All patients began tamoxifen within 6 days of surgery and continued this for 5 years. Kaplan-Meier methods, the log-rank test, and multivariable Cox regression models were used to assess differences in five-year disease-free survival (DFS) between the groups. All statistical tests were two-sided.

Results: The randomized midluteal phase surgery group had a five-year DFS of 64%, compared with 71% for the immediate surgery random assignment group (hazard ratio [HR] = 1.24, 95% confidence interval [CI] = 0.91 to 1.68, P = .18). Multivariable Cox regression models, which included important prognostic variables, gave similar results (aHR = 1.28, 95% CI = 0.94 to 1.76, P = .12). For overall survival, the univariate hazard ratio was 1.33 (95% CI = 0.94 to 1.89, P = .11) and the multivariable aHR was 1.43 (95% CI = 1.00 to 2.06, P = .05). Better DFS for follicular phase surgery, which was unanticipated, proved consistent across multiple exploratory analyses.

Conclusions: The hypothesized benefit of adjuvant luteal phase oophorectomy was not shown in this large trial.

Over a century ago, Beatson reported on the favorable impact of surgical oophorectomy in a young woman with metastatic breast cancer who had a four-year remission of her disease (1,2). While other surgeons subsequently demonstrated remissions in patients with breast cancer with surgical oophorectomy, the high rate of mortality from this procedure in the early 20th century discouraged many (3,4). In 1992, the Early Breast Cancer Trialists' Collaborative Group published meta-analysis data demonstrating benefits from adjuvant oophorectomy by radiation or surgery and in subsequent years mostly medical oophorectomy (with gonadotropin-releasing hormone agonists) has been extensively evaluated and considered equivalent as adjuvant therapy (5). In 2002, some of the current report authors published a communication on an exploratory post hoc analysis of subsets of participants in a clinical trial which suggested that adjuvant surgical oophorectomy accomplished in the historical luteal phase of the menstrual cycle was more effective than if this surgery was done in the follicular phase (6). An editorial discussing these results suggested that only a prospective randomized trial could assist in resolving the veracity of this observation (7).

In this communication we report the mature results of a phase III randomized clinical trial of the timing of surgical oophorectomy in the treatment of premenopausal women with hormone receptor–positive breast cancer.

Methods

Design and Eligibility

Between April, 2003 and October, 2009, at clinical sites in the Philippines (#6), in Vietnam (#3), and in Morocco (#1), we recruited 740 premenopausal women with core biopsy hormone receptor–positive (estrogen receptor– or progesterone receptor–positive in experienced local laboratories by immuno-histochemical assays) invasive and operable breast cancers where all subjects were treated with surgical oophorectomy and mastectomy (in that order) under the same anesthesia, followed by tamoxifen for five years. At registration 509 women (69%), based on their menstrual histories, would not be in the luteal phase of their menstrual cycle in the following one to six days. These women were randomly assigned to undergo these surgeries either immediately in the next one to six days or delayed until a five-day window of time during the midluteal phase in a future-estimated menstrual cycle (randomly assigned patients).

Another 231 women (31%) were identified to be in the luteal phase of their menstrual cycle in the following one to six days, based on their menstrual histories. These women were scheduled to undergo immediate surgical therapies with surgical oophorectomy and mastectomy during that next one-to-six-day period (nonrandomized patients). This design was directed to answering the specific question that patients and their treating physicians had about the luteal phase oophorectomy benefit hypothesis: "If the patient is not currently in her luteal phase when she would have her surgeries, should her surgeries be delayed until the next luteal phase?" This design minimized the numbers of patients with delayed surgeries to those only who were surmised to potentially benefit. We recognized that this design addressed a clinical question about optimal timing, rather than more directly the follicular phase–luteal phase oophorectomy surgery relative benefits. Other designs that involve hormonal assays before random assignment were considered and deemed unfeasible. Thus the randomized patients were on historical day 0 to 13 or day 23 to 35 of their menstrual cycle at registration, while nonrandomized patients were on day 14 to 22 of their cycle at registration. The time window between biopsy and registration on study was approximately two weeks. Random assignment was implemented through a central coordinating center in the United States. Computer-generated random assignment was in a 1:1 ratio for each surgery protocol, was blocked in allocation, and was stratified by clinical site.

Eligibility criteria for the trial included: age 18 to 50 years (inclusive), history of menstrual cycles for last three months of 25 to 35 days (inclusive) and last menstrual period less than 35 days ago, not taking oral contraceptives, pathologic histologic diagnosis of invasive and hormone receptor–positive breast cancer by local IHC analyses, clinical stage II to IIIB, physical examination including gynecological examination unrevealing for any suggestion of serious illness or metastatic breast cancer, chest x-ray unrevealing for evidence of serious pulmonary disease (such as tuberculosis) or metastases, and negative urine pregnancy test. Phenotype Luminal A cases are those defined as ER+PR+Her-2/neu-.

The study was approved at individual participating institutions in the Philippines, Vietnam, and Morocco and/or by supervising institutional review boards (IRBs) for these institutions and at the lead investigator's American institutions. Annual written reapprovals by all IRBs were also obtained. All participants provided written informed consent. All of the participating institutions and IRBs were registered with the Office for Human Subjects' Protection (OHRP) in the United States. A data and safety monitoring committee of six individuals approved by the United States National Institutes of Health reviewed the trial conduct, primary, secondary, and safety results annually. The trial was registered at ClinicalTrials.gov number, NCT 00201851, on September 12, 2005. Further details of design and recruitment are given in Supplementary Methods (available online).

Statistical Methods

Sample Size
At the initial planning of this study, information from exploratory analyses of previous trial data was available with median follow-up of approximately 3.5 years (6). At that time it was estimated that five-year disease-free survival (DFS) for patients receiving scheduled (midluteal phase) surgery would fall between 80% and 84%, while those patients having immediate surgery (follicular phase) would have five-year DFS probabilities between 64% and 72%. With 340 randomly assigned participants (a total sample size of 510), at least 80% power was expected to detect a net difference of approximately 14% (scheduled surgery group five-year DFS of 80% to 84% and immediate surgery group five-year DFS of 66% to 70%) by a log-rank test and a two-sided 5% type I error rate.

Reassessment of Sample Size
Four years after the original planning for this study, analysis of the benefit of luteal vs follicular phase surgery was reevaluated using new long-term follow-up information. Three important issues were identified. First, the overall failure (recurrence) rate in the new current study was lower than had been originally projected. Second, in the data from the hypothesis-generating study, the observed effect size decreased with longer-term follow-up (8). Third, the hazard rates for recurrence were observed to converge after 5.5 years. Given the additional information from the current trial, we reestimated sample size with blinding intact and proposed an increase in the originally planned

recruitment number from 340 to 510 randomly assigned (762 total) (9). We expected that this new estimate of sample size would have at least 80% power to detect a hazard ratio in the range of 0.62 to 0.65.

Analyses

DFS and overall survival (OS) curves were calculated using Kaplan-Meier methods, and the differences in survival curves were assessed via the log-rank test. The planned primary hypothesis test compared DFS between the randomized groups by the log-rank test. The Cox proportional hazards model was used to estimate univariable hazard ratios, and multivariable Cox models were used to estimate adjusted hazard ratios (aHRs). Proportionality assumptions for the Cox models were assessed by diagnostic plots of the scaled Schoenfeld residuals and log-minus-log survival plots. Substantial deviations from proportionality were not observed. DFS was defined as the time from the date of random assignment to the date of first recurrence. Patients without recurrence were censored at the date of their last follow-up where they were known to be disease-free. OS was defined as the time from random assignment to the date of death or to their censoring date for last follow-up when they were known to be alive. In all comparisons of the randomized patients, treatment group assigned at random assignment was compared, regardless of the treatment received. Exploratory analyses were used to confirm or explain findings. P values in exploratory analyses were reported for completeness, were not corrected for multiplicity, and cannot be interpreted as formal hypothesis tests.

Results

After 740 patients had entered the study, we believed that we reached our goal with 511 randomized patients, and patient recruitment was discontinued. Subsequently on audit two of these randomly assigned cases were found to be duplicates, and thus the final CONSORT diagram lists 509 randomly assigned case patients. The number of patients in each randomly assigned treatment group and in the non–randomly assigned group who received the intended therapy, who were ineligible, or who refused therapy are summarized in Figure 1. Table 1 provides detailed patient and disease characteristics by treatment (randomly assigned and non–randomly assigned) groups. Group labels are simplified to A for those randomly assigned to delayed luteal phase surgery, B for those randomly assigned to immediate surgery, and C for those not randomly assigned. Of note is that 91%, 96%, and 94% of patients in Groups A, B, and C received treatment in their assigned window. Among all the patients who did not have surgery in the assigned time windows, the majority of cases were for reasons of a compelling and rational medical or psychosocial nature. The progesterone levels were higher in Groups A and C than in Group B, as expected. Group A appeared more advanced than Groups B and C in stage, percentage with positive axillary nodes, and percentage with four or more positive axillary nodes. Thirteen patients distributed across groups and sites received some form of chemotherapy.

The current report is based on evaluations to May 1, 2013. Maximum follow-up for disease-free survival had been set a priori at six years (9). At this time point, the median follow-up was 5.0 years.

Table 2 describes when patients in the three groups actually had their surgeries, according to their histories. For Group B, 66% actually had their surgeries during follicular phase. This was the percentage expected during the design stage, and these data further confirm that the timings of surgeries were predominantly as assigned. No unusual features were observed in terms of the date of registration for Groups A and B with respect to day of the patient's menstrual cycle.

Although a much higher proportion of patients in Groups A and C had high progesterone values (5 ng/mL or greater) than
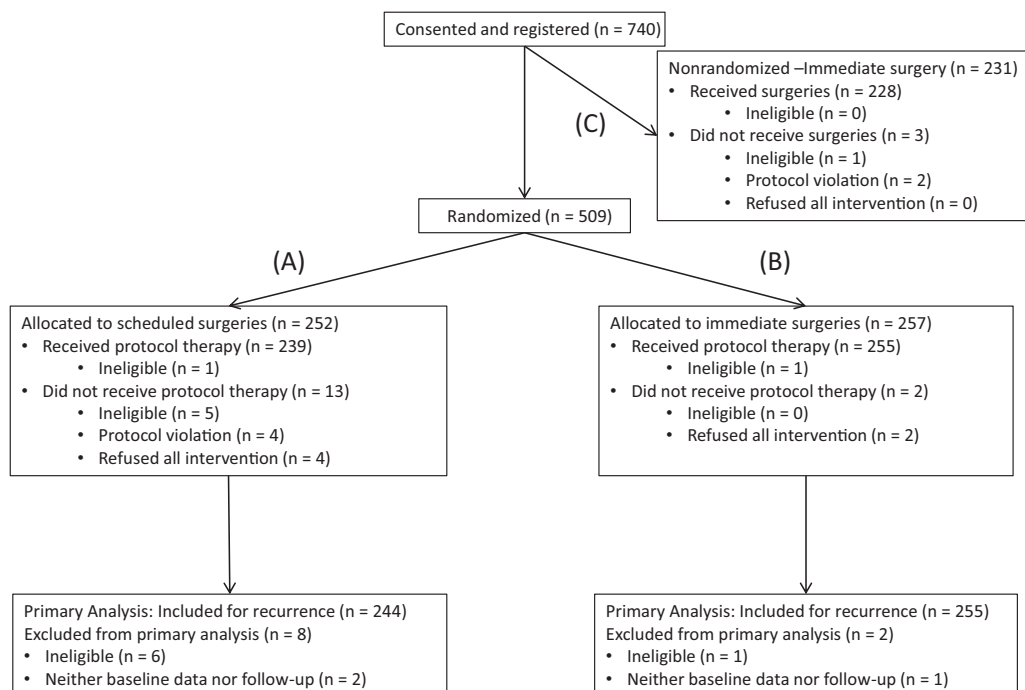

Figure 1. CONSORT diagram for the trial. Reasons for ineligibility included: core biopsy tissue hormonal receptor negative for both estrogen receptor and progesterone receptor; on review, no evidence of invasive cancer; pregnancy; thyroid toxicity; and taking oral contraceptives.

those in Group B, there were still many with low or moderate levels (Table 1). Only slightly more than half of all luteal phase history patients actually had progesterone levels at 5 ng/mL or higher, confirming biological ovulatory event and secretory phase. One-third of cases in each of Groups A and C, despite histories of luteal phase, had progesterone levels under 2 ng/mL, consistent with follicular phase or an anovulatory cycle. This frequency of anovulatory cycles is consistent with the data showing that 44% of the study population was over age 44 years (Table 1), the stressful circumstances of the women just diagnosed with cancer, and menstrual bleeding in perimenopausal, often multiparous women interpreted as regular cyclical menstruation. Furthermore, large variation in progesterone levels was observed in all treatment groups (data not shown).

To further explore the relationship between progesterone and cycle, we examined those levels based on history of menstrual cycle. As Table 3 shows, of all patients with a history of follicular phase surgeries and with progesterone levels (n = 167), 75% had progesterone levels under 2 ng/mL, and 86% had progesterone levels under 5 ng/mL consistent with that history. In contrast to the 500 patients whose surgeries were done more than 14 days from LMP and with progesterone levels, 182 patients (36%) had progesterone levels of under 2 ng/mL. Among these patients with estradiol data, estradiol levels were at 100 pg/mL or higher in 83 of 181 (46%), consistent with an anovulatory cycle. Estradiol levels were under 100 pg/mL in 98 of 181 (54%), which is more consistent with noncycling status.

Completeness of Follow-Up Information

Overall, 714 of the 729 primary analysis set patients (97.9%) either had follow-up out to five years or to death or had a

Table 1. Frequencies and column percentages of patient and disease characteristics by assigned group*

Variable	Level	A: Scheduled (n = 244) No. (%)	B: Immediate (n = 255) No. (%)	C: Nonrandomized (n = 230) No. (%)
Age, y	< 44	135 (55.3)	140 (54.9)	131 (57.0)
	≥ 44	109 (44.7)	115 (45.1)	99 (43.0)
Core biopsy ER hormone receptor assay results	Positive	226 (93.4)	235 (92.5)	217 (94.3)
	Negative	16 (6.6)	19 (7.5)	13 (5.7)
Mastectomy ER results	Positive	176 (93.6)	188 (91.7)	174 (94.1)
	Negative	12 (6.4)	17 (8.3)	11 (5.9)
Core biopsy PR hormone receptor assay results	Positive	218 (89.7)	234 (92.1)	219 (95.2)
	Negative	25 (10.3)	20 (7.9)	11 (4.8)
Mastectomy PR results	Positive	163 (86.7)	171 (83.4)	165 (88.2)
	Negative	25 (13.3)	34 (16.6)	22 (11.8)
Mastectomy HER2 result (Herceptest)	Positive	41 (21.9)	45 (22.0)	32 (17.3)
	Negative	146 (78.1)	160 (78.0)	153 (82.7)
Combined site and reference pathologist histologic grade	1	39 (18.8)	42 (18.8)	35 (16.7)
	2	117 (56.3)	126 (56.5)	133 (63.6)
	3	52 (25.0)	55 (24.7)	41 (19.6)
Surgical treatment received in protocol prescribed time window	Received in window	223 (91.4)	244 (95.7)	217 (94.3)
	Received outside of window	15 (6.1)	10 (3.9)	11 (4.8)
	No surgery received	6 (2.5)	1 (0.4)	2 (0.9)
Pathological stage	I-II	136 (56.2)	162 (63.8)	143 (62.2)
	III-IV	106 (43.8)	92 (36.2)	87 (37.8)
Pathologic tumor size, cm	Mean (SD) (range)	4.5 (2.8) (1.0 – 22.0)	4.7 (3.2) (0.8 – 20.3)	4.1 (2.3) (0.8 – 15.0)
Number of positive axillary nodes	None	96 (40.5)	111 (43.7)	103 (45.4)
	1–3	60 (25.3)	73 (28.7)	56 (24.7)
	4+	81 (34.2)	70 (27.6)	68 (30.0)
Adjuvant radiotherapy given after surgical mastectomy	No	146 (60.1)	154 (60.9)	139 (60.4)
	Yes	97 (39.9)	99 (39.1)	91 (39.6)
Progesterone level on day of surgery, ng/mL	<2	75 (34.1)	158 (66.7)	74 (34.4)
	2- <5	27 (12.3)	22 (9.3)	24 (11.2)
	≥5	118 (53.6)	57 (24.1)	117 (54.4)

* ER = estrogen receptor; HER2 = human epidermal growth factor receptor 2 (ERBB2); PR = progesterone receptor.

Table 2. Frequencies and column percentages of menstrual cycle phase on day of surgery estimated by history by assigned group

	Group			
	A	B	C	
Phase on day of surgery	No. (%)	No. (%)	No. (%)	Total
Follicular	9 (3.8)	166 (65.9)	8 (3.5)	183
Luteal	229 (96.2)	86 (34.1)	222 (96.5)	537
Total	238	252	230	720

follow-up visit within the six months before May 1, 2013. These percentages for Groups A, B, and C were 96.3%, 98.4%, and 99.1%, respectively.

Adverse Events

No patients died in the 30 days following surgeries, and four developed major morbidities: pneumonia (2) and venous thrombosis (2). Adverse events that were clearly treatment related included: 1) pregnancy (normal infant born after exposure to four months of tamoxifen); 2) endocervical cancer (1:3490 patient years); and 3) six uterine polyps.

Symptoms associated with oophorectomy and tamoxifen were limited and very similar to those previously reported for this treatment among Vietnamese women (10). No patients acknowledged stopping tamoxifen because of symptomatic side effects.

Primary Hypothesis Testing

Using data as defined in the CONSORT diagram (Figure 1) of all entered and randomly assigned patients, five-year DFS was 64% in the scheduled luteal phase surgery group (A), compared with 71% for the immediate surgery randomized group (B) (log-rank $P = .18$, hazard ratio [HR] = 1.24, 95% confidence interval [CI] = 0.91 to 1.68) (Figure 2). To be clear, this nonsignificant log-rank test result and hazard ratio were actually in the opposite direction of expected.

We also compared Groups A and B using a multivariable Cox model that included stage, pathologic tumor size, adjuvant radiation therapy, axillary nodal status, and age, covariates that have been identified previously to influence survival. All of these prognostic variables showed independent associations with DFS. In multivariable Cox regression analyses including these variables, the adjusted hazard ratio (1.28, 95% CI = 0.94 to 1.76, $P = .12$) changed little from the univariate hazard ratio (Table 4).

The pattern for OS for Groups A and B was similar to that observed for DFS (Figure 3). The HR for death comparing A and B was 1.33 (95% CI = 0.94 to 1.89, log-rank $P = .11$) and increased to 1.43 (95% CI = 1.00 to 2.06, $P = .05$) after adjustment.

Exploratory Analyses

Because of the unexpected worse survival in Group A, we explored DFS in Group C, the non–randomly assigned immediate surgery luteal phase group. The Kaplan-Meier curve for Group C looked similar to that of Group B (Immediate Follicular), and so deviated similarly from Group A. A multivariable Cox model that compared DFS of Groups A and C showed an aHR of 1.19 (95% CI = 0.86 to 1.65, $P = .30$).

Because one designed difference between Groups A vs B and C was that Group A patients experienced delays in surgery until they were in their historical midluteal phase, we explored the relationship between length of delay and DFS among Group A patients only. Patients in Group A varied as a

Table 3. Frequencies and column percentages of progesterone level by menstrual cycle phase on day of surgery estimated by history

Progesterone on day of surgery, ng/mL	Phase on day of surgery		
	Follicular No. (%)	Luteal No. (%)	Total
<2	125 (74.9)	182 (36.4)	307
2-<5	18 (10.8)	53 (10.6)	71
≥5	24 (14.4)	265 (53.0)	289
Total	167	500	667

Number at risk

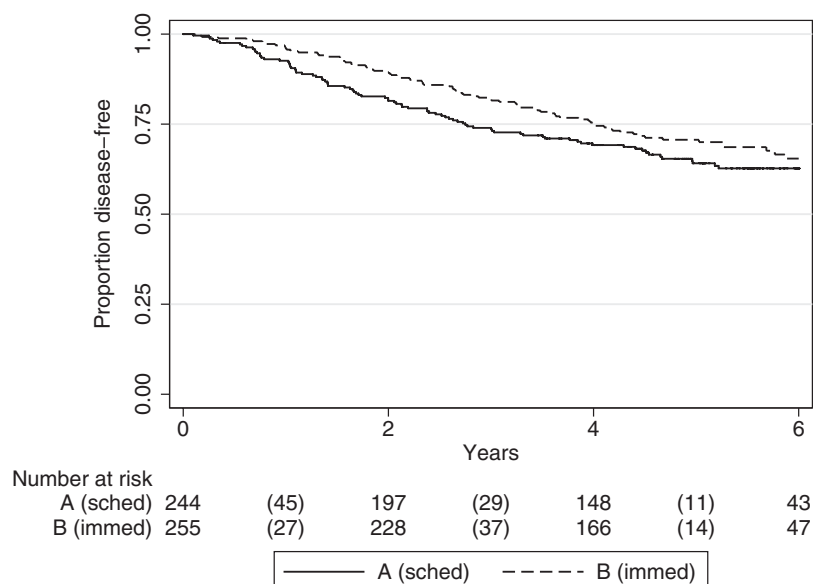|						
A (sched)	244	(45)	197	(29)	148	(11)	43
B (immed)	255	(27)	228	(37)	166	(14)	47

——— A (sched) - - - - - B (immed)

Figure 2. Disease-free survival for randomized groups. Kaplan-Meier disease-free survival curves for the randomized Groups A (scheduled surgery) and B (immediate surgery) (two-sided log-rank $P = .18$). Number of events is indicated in parentheses.

function of the timing protocol, half were delayed less than two weeks, and 25% were delayed three weeks or more. The length of delay was divided into quartiles for inclusion in a multivariable model. Several patients experienced extremely long delay times because of non–study related factors, and categorizing delay time avoided lending an undue influence to these large values. Analysis by quartiles showed for 11 to 14 vs 0 to 10 days' delay an aHR of 0.78, (95% CI = 0.39 to 1.54, P = .47); for 15 to 19 vs 0 to 10 days' delay an aHR of 1.35 (95% CI = 0.67 to 2.70, P = .40); and 20+ vs 0 to 10 days' delay aHR of 1.15 (95% CI = 0.59 to 2.22, P = .68).

DFS by axillary nodal status showed that in axillary node negative patients Group A vs Group B: aHR of 0.72 (95% CI = 0.34 to 1.53, P = .39); in axillary node positive patients Group A vs B: an HR of 1.48 (95% CI = 1.05 to 2.10, P = .027). Of 229 recurrences (A, B, and C groups), 60 (26.2%) were local only (defined as recurrence in chest wall, axillary nodes, or supraclavicular nodes on the side of the original cancer without additional sites indicated). Of these 60 local-only recurrence cases, 15 (25.0%) had received adjuvant radiotherapy.

Our hormonal studies clearly demonstrated that our patients' historical menstrual cycle information incompletely or inaccurately identified patients as being in regular cycling follicular or luteal phase. In the randomized groups, Group A patients with higher Pg (≥2ng/mL) had better DFS than those with less than

Table 4. Multivariable DFS Cox model results for randomized groups (n = 486)*

Variable	HR (95% CI)	P
A vs B	1.28 (0.94 to 1.76)	.1180
Adjuvant radiotherapy vs none	0.69 (0.49 to 0.97)	.0320
Stage III/IV vs I/II	2.79 (1.58 to 4.92)	.0004
1–3 positive nodes vs none	1.86 (1.13 to 3.06)	.0151
4+ positive nodes vs none	2.27 (1.23 to 4.19)	.0087
Log of max pathologic size	1.39 (1.02 to 1.90)	.0390
Age (1-year increase)	0.94 (0.91 to 0.97)	.0002

* P values calculated from two-sided Wald tests. DFS = disease-free survival.

2ng/mL (aHR = 0.53, 95% CI = 0.34 to 0.84, P = .006). When hormonal and historical information are considered together, three distinct groups are defined: 1) Patients with a history consistent with follicular phase and progesterone levels of under 2ng/mL ("confirmed follicular"). These are all not necessarily regular cycling patients. 2) Patients with a history (>14 days from LMP) consistent with luteal phase and a progesterone level of 5ng/mL or higher, suggesting a normal cycling luteal/secretory phase ("confirmed luteal"). 3) Patients with a history (>14 days from LMP) consistent with luteal phase, but progesterone level of under 2ng/mL, suggesting an anovulatory cycle ("unconfirmed luteal"). In a multivariable analysis including all study patients (Groups A, B, and C), weak trends for better DFS in Group I ("confirmed follicular") vs Group II ("confirmed luteal") patients and worse DFS in Group III ("unconfirmed luteal") vs Group II ("confirmed luteal") were observed (Figure 4 and Table 5). Other subgroup analyses gave similar results. For example, after restricting analysis to Group B patients only, those in the confirmed follicular group showed better DFS than those in the confirmed luteal group (aHR = 0.74, 95% CI = 0.40 to 1.40, P = .36). This result is also of interest, because there is no difference between the delay in surgery between these defined follicular and luteal groups.

Finally, irrespective of timing of surgical oophorectomy, the combined hormonal treatment evaluated in this study was associated with excellent outcomes in phenotype Luminal A axillary node negative patients, with a DFS of 89% (95% CI = 82–93%) and an OS of 95% (95% CI = 89% to 97%) at five years of follow-up.

Discussion

The detailed data provided in this report support a conclusion that this trial was properly conducted as planned. In total this report suggests that multicenter international trials, including non–high-income countries, can be successfully pursued. This prospective randomized study was specifically designed to test the hypothesis that surgical oophorectomy in the historical luteal phase of the menstrual cycle is associated with better outcomes. However, the hormonal studies on the majority of study participants here indicate that menstrual cycle history
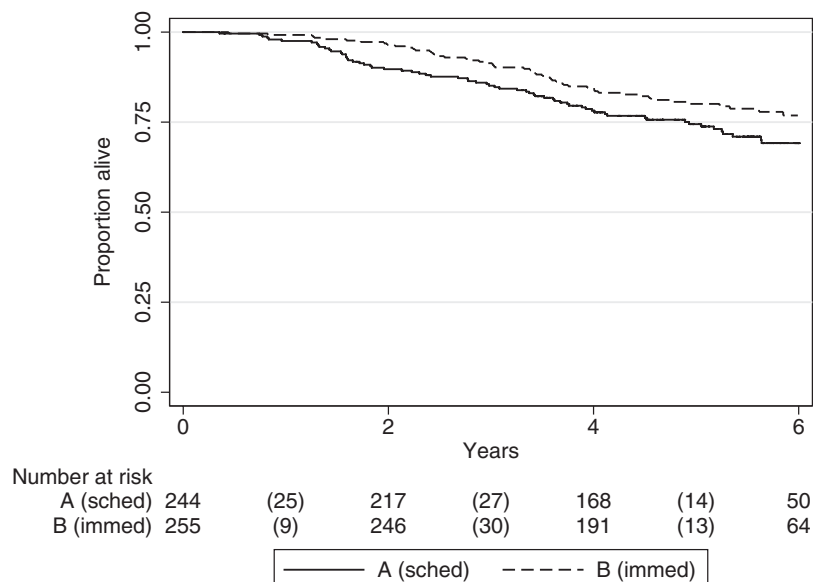

Figure 3. Overall survival for randomized groups. Kaplan-Meier overall survival curves for the randomized Groups A (scheduled surgery) and B (immediate surgery) (two-sided log-rank P = .11). Number of events is indicated in parentheses.
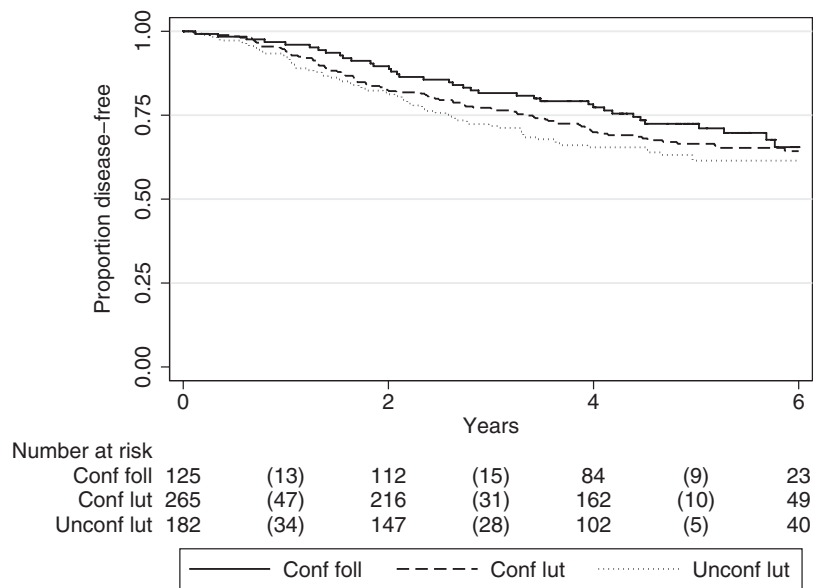
Figure 4. Disease-free survival for confirmed follicular, confirmed luteal, and unconfirmed luteal groups. Kaplan-Meier disease-free survival curves for confirmed follicular (conf foll), confirmed luteal (conf lut), and unconfirmed luteal (unconf lut), including randomized and nonrandomized patients. Number of events is indicated in parentheses.

Table 5. Multivariable Cox models results for DFS comparing confirmed follicular, confirmed luteal, and unconfirmed luteal including randomized and nonrandomized patients (n = 569)*

Variable	HR 95% CI	P
Conf foll vs conf lut	0.82 (0.56 to 1.21)	.1739
Unconf lut vs conf lut	1.20 (0.87 to 1.66)	
Adjuvant radiotherapy vs none	0.59 (0.43 to 0.82)	.0016
Stage III/IV vs I/II	2.19 (1.28 to 3.73)	.0042
1–3 positive nodes vs none	2.39 (1.51 to 3.79)	.0002
4+ positive nodes vs none	2.98 (1.64 to 5.41)	.0003
Log of max pathologic size	1.30 (0.95 to 1.76)	.0977
Age (1-year increase)	0.95 (0.92 to 0.98)	.0010

* P values calculated from two-sided Wald tests. DFS = disease-free survival.

alone offers an unreliable assessment of biological menstrual cyclicity. The percentage of anovulatory patients identified in this study is similar to those identified in an NCCTG breast cancer surgery timing study, when the different cutoff definitions are taken into account ([11]). Exclusion, ineligibility, missing data, compliance with assigned treatments, follow-up, and fraction of case patients with blood studies numbers are all of magnitudes desired; the central pathology laboratory assessments of critical ER, PR, Her-2/neu, and histologic tumor grade are reliable. The study results are not complicated by bisphosphonate treatments.

The sample size for this trial was large enough to obtain a significant DFS hazard ratio for moderately superior luteal scheduled surgeries (80% power for HR = 0.64). Given this power and the fact that the actual primary results were opposite of what was expected and showed a sufficiently narrow confidence interval allows the tentative claim that delaying surgeries until luteal phase provides no substantial benefit. A major explanation for the negative result was that a high proportion of women assigned to luteal surgery had low progesterone levels. Specifically, in this study the relationships of menstrual cycle timing and oophorectomy surgery appear affected by three circumstances: 1) the design of the study led to only 121 out of 457 randomly assigned patients with known hormonal levels at surgery, identifiable as historically in follicular phase and with

progesterone levels under 2 ng/mL, consistent with that history ([Table 3]). 2) Close to half of patients with a history of luteal phase surgery had progesterone levels under 5 ng/mL, inconsistent with being in a regular ovulatory menstrual cycle in luteal phase ([Table 3]). Most of these patients must have been in anovulatory cycles or not having menstrual cycling at all. 3) The worst DFS was found for women who had surgeries done in historical luteal phase associated with progesterone levels of under 2 ng/mL ([Figure 4] and [Table 5]).

Further explicatory biological mechanism hypotheses are complicated by the understandings that even with specific hormonal assessments, we can only approximate pictures of the circumstances at the precise time of definitive oophorectomy and breast surgeries, and other preceding hormonal events and perioperative surgical events (increased tumor circulating cells for example) may be playing substantial cell-signaling roles. That surgical oophorectomy should have immediate (and not only chronic) cell signaling effects (and thus should give supplementary benefit beyond that of tamoxifen alone) should not be surprising given the prompt salutary effects of orchiectomy in men with metastatic prostate cancer. The general findings here regarding the inaccuracies of defining luteal phase menstrual cyclicity also may partially explain the apparently misleading findings from the hypothesis-generating study ([6]).

While scientific uncertainty may characterize the results offered here, there would seem to be some conclusions for clinical practice. Menstrual cycle history is an increasingly inaccurate tool for describing biological hormonal cyclicity as women in their 40s. Given the high frequency of perimenopausal status found by hormone assessments among the patients in this and other studies, progesterone level determination may be helpful in planning therapy ([11]). Finally, the data here on outcomes for axillary node–negative phenotype Luminal A, clinical stage II to IIA treated with adjuvant surgical oophorectomy and tamoxifen support serious consideration of this approach alone as an option for all women in this favorable subgroup ([12], [13]).

As planned, this study does not look at outcomes after five years, because no treatments were given after five years, and the hazard function for recurrence, consistent with the results of randomized studies of post–five-year treatment, rises at that

8 of 8 | *JNCI J Natl Cancer Inst*, 2015, Vol. 107, No. 6

time (9, 14–16). Based on recent data, our study patients might have benefited from continuation of their tamoxifen therapy after five years (17). In this study, potentially beneficial adjuvant radiation therapy was given in approximately 40% of case patients, when by high-income country treatment guidelines this intervention was indicated in 80% (18).This omission did not influence the study results, and the distribution of this missed radiation treatment was similar in the two randomly assigned and in the non–randomly assigneded patients (Table 1). Exploratory analyses in subgroups included radiation therapy as a covariate for adjustment. In sum, better outcomes associated with surgical oophorectomy plus tamoxifen treatment should be expected with longer-term tamoxifen treatment and adjuvant radiation therapy for all higher-risk cases.

An ancillary study of a subset of patients in the current report has shown that there was no loss of bone mineral density (BMD) in femoral neck and stabilized loss in LS spine with surgical oophorectomy and tamoxifen treatment, indicating no therapeutic need for bisphosphonate for BMD preservation with surgical oophorectomy and tamoxifen treatment (19).

The data here support future investigation of surgical oophorectomy plus tamoxifen in hormonally confirmed cycling premenopausal women vs defined standard optimal hormonal therapy or therapies emerging from current studies, when as for phenotype Luminal A cases, hormonal therapy alone appears justified (12,13). Further, investigation of this hormonal strategy with and without chemotherapy in higher risk patients may be considered. Finally, the better characterization of at-risk women by hormonal testing prior to treatment with prophylactic oophorectomy for breast cancer warrants investigation.

In premenopausal women with operable breast cancer that is hormone receptor–positive, adjuvant treatment with surgical oophorectomy that is in the historical luteal phase was not shown to be more beneficial than surgery performed at other times. Exploratory analyses suggested that this timing of surgery may be less effective, in part because of the fact that in this study almost half of women in historical luteal phase appeared by hormonal studies to be having anovulatory cycles or were perimenopausal.

Funding

This study was funded by the United States National Institutes of Health (RO1 CA 097375), the Breast Cancer Research Foundation, and the International Breast Cancer Research Foundation. The sponsors had no roles in the implementation, conduct, analysis, or interpretation of the study after agreeing to provide the funding.

Notes

Author contributions: R. Love and S. Anderson designed the initial study proposal. E. M. Hade, G. S. Young, and D. Jarjoura developed the sample size modification analysis. R. Love developed the database and managed the study and data collection, site investigator recruitment and training, and regulatory issues, and wrote the study report. A. Laudico, N. V. Dinh, G. Uy, L. H. Quang, J. Salvador, S. S. Siguan, R. Mirsol-Lumague, N. D. Tung, N. Benjaafar, N. Navarro, T. Quy, A. DeLa Pena, R. Dofitas, O. Bisquera, and N. Linh recruited and managed the study patients and collected and submitted the study data. T. To conducted pathology studies and managed the tissue samples from the Vietnamese sites. D. C. Allred conducted the central laboratory mastectomy specimen estrogen receptor, progesterone receptor, Her-2/neu, and histological grading studies for all study cases. G. S. Young, E. M. Hade, and D. Jarjoura collected and managed the study data, conducted the statistical analyses, and contributed to the writing of this manuscript and its preliminary reports. All authors concurred with the report submitted.
The authors declare that they have no conflicts of interest.

The data safety and monitoring committee for this trial: Nancy Davidson, David Mahvi (Ch), Kathleen Pritchard, David Rimm, Heidi Sahel, and Barry Storer.
Other acknowledgements: Lou Ann Stittleburg, Barbara Braunger, Linda McCart, Dorsia Wakawa, Heather Story, Ruth Nituda, Susan Anderson, Ophira Ginsburg, Susan Carlson, Seema Kahn.

References

1. Beatson CT. On treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment with illustrative cases. *Lancet*. 1896;2:104–107.
2. Thomson A. Analysis of cases in which oophorectomy was performed for inoperable carcinoma of the breast. *BMJ*. 1902;2:1538–1541.
3. Boyd S. On oophorectomy in cancer of the breast. *BMJ*. 1900;2:1161–1167.
4. Love RR, Philips J. Oophorectomy for breast cancer: history revisited. *J Natl Cancer Inst*. 2002;94(19):1433–1434.
5. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. Early Breast Cancer Trialists' Collaborative Group. *Lancet*. 1992;339(8785):71–85.
6. Love RR, Duc NB, Dinh NV, et al. Mastectomy and oophorectomy by menstrual cycle phase in women with operable breast cancer. *J Natl Cancer Inst*. 2002;94(9):662–669.
7. Hortobagyi GN. The Influence of Menstrual Cycle Phase on Surgical Treatment of Primary Breast Cancer: Have We Made Any Progress Over the Past 13 Years? *J Natl Cancer Inst*. 2002;94(9):641–643.
8. Love RR, Young GS, Hade EM, et al. Effects on survival of menstrual cycle phase of adjuvant surgical oophorectomy in premenopausal women with breast cancer. *Breast Cancer Res Treat*. 2011;126(2):479–485.
9. Hade EM, Jarjoura D, Lai W. Sample size re-estimation in a breast cancer trial. *Clin Trials*. 2010;7(3):219–226.
10. Love RR, Nguyen BD, Nguyen CB, et al. Symptoms associated with oophorectomy and tamoxifen treatment for breast cancer in premenopausal Vietnamese women. *Breast Cancer Res Treat*. 1999;58(3):281–286.
11. Grant CS, Ingle JN, Suman VJ, et al. Menstrual cycle and surgical treatment of breast cancer: findings from the NCCTG N9431 study. *J Clin Oncol*. 2009;27(22):3620–3626.
12. Coates AS, Colleoni M, Goldhirsch A. Is adjuvant chemotherapy useful for women with luminal a breast cancer? *J Clin Oncol*. 2012;30(12):1260–1263.
13. Hayes DF. Targeting adjuvant chemotherapy: a good idea that needs to be proven! *J Clin Oncol*. 2012;30(12):1264–1267.
14. Goss PE, Ingle JN, Martino S, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *New Engl J Med*. 2003;349(19):1793–1802.
15. Ingle JN, Tu D, Pater JL, et al. Duration of letrozole treatment and outcomes in the placebo-controlled NCIC CTG MA.17 extended adjuvant therapy trial. *Breast Cancer Res Treat*. 2006;99(3):295–300.
16. Love RR, Van Dinh N, Quy TT, et al. Survival After Adjuvant Oophorectomy and Tamoxifen in Operable Breast Cancer in Premenopausal Women. *J Clin Oncol*. 2008;26(2):253–257.
17. Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years vs stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet*. 2013;381(9869):805–816.
18. NCCN.com. Treatment guidelines for breast cancer. Sept 17, 2013.
19. Love RR, Young GS, Laudico AV, et al. Bone mineral density following surgical oophorectomy and tamoxifen adjuvant therapy for breast cancer. *Cancer*. 2013;119(21):3746–3752.

ARTICLE