Reviewer Comment:

1.The participant population was predominantly white and well-educated. [Information about educational status was not included in the electronic database.]. Few non-white women were included in the study; the small numbers of women from African-American and Other racial groups were balanced between treatment arms.

2. Sixty-three percent of women in the trial had an intact uterus and were therefore at risk for tamoxifen-induced endometrial cancer at the start of the study. An additional 213 women on placebo (3% of study arm) and 334 on tamoxifen (5%) had a hysterectomy during the course of the study. These figures exclude participants who were diagnosed with uterine cancer. The reason for this unbalance is not known.

3. Approximately 1/3 of the participant population was premenopausal (intact uterus and ovaries with continued menstrual periods). An additional 2955 participants (1450 on placebo, 1505 on tamoxifen) had undergone hysterectomy without an oophorectomy; menopausal status in these women was unknown.

4. The median weight of participants was 154 lbs. The range of weights was broad, from 100 to 382 lbs. Weight is a risk factor for endometrial cancer, deep vein thrombosis, and cardiac disease. Weight ranges were equally distributed between treatment arms, and weight ranges appeared comparable across prospectively defined age groups.

5. Cardiovascular risk factors and smoking history are discussed in Section 10.4, Cardiovascular events.

6. As shown in Table 9, the electronic database was used to evaluate the comparability of the two treatment groups. Breast cancer risk factors, including the prospectively defined age groups, number of first degree relatives with breast cancer (0, 1, or 2 or more), history of LCIS, history of atypical hyperplasia, number of breast biopsies (0, 1, or ≥ 2), age at first live birth (no live births, 12-19, 20-24, 25-29, and 30 or more), and age at menarche (≥ 14, 12-13, ≤ 11), were well-balanced between treatment groups. Family history by age and history of LCIS by age were also balanced.

7. Approximately half the women in the trial had an average age of menarche (12-13) and no history of a previous breast biopsy.

8. Most women in the trial had children, and nearly all had given birth before the age of 30. Because the age at first live birth has increased in the United States, a greater number of younger women may meet the BCPT entry criteria in the future. This possibility has implications for the recruitment to P-2, the STAR trial, as well as for use of tamoxifen in the general population.

9. While most women did not have a history of atypical hyperplasia or LCIS, the largest number of women with these pathologic entities ever followed in a controlled trial was entered on P-1.

10. Seventy-five percent of the women in the trial had at least one first-degree relative with breast cancer. Conversely, approximately 80% of the P-1 population had no family history of breast cancer or a history in one first-degree relative only. These data suggest that the number of women with a history of breast cancer linked to BRCA1 and/or BRCA2 mutations may be low. Evaluation of BRCA1/2 status in the P-1 population will be of great interest.

11. The risk categories listed in Table 9 by the NSABP differ from those prospectively defined as stratification factors (see Section 8.2.2, Trial entry).

12. We received the Gail model software 8/14/98, too late to allow us to independently verify risk levels. Because we have some risk factors by grouping rather than as an actual data point (menarche grouped as described above rather than as a specific age per participant), we cannot replicate entry levels of risk.

8.7 On-study therapy

8.7.1 Compliance

Compliance was evaluated on a prospectively defined schedule. Participants were to be contacted at 1 month, 3 months, 6 months, and then every 6 months to monitor and promote compliance. Participants were asked to provide self-assessments of compliance. Pill counts were also performed at each follow-up visit by the study coordinator. Forms denoting these contacts were to be submitted on a regular basis to the NSABP. Patients who withdrew from study were asked to undergo an exit interview with the on-site protocol coordinator, and an Off Therapy Form was submitted. This form included a space in which to document the exit interview, and to note what interventions were provided by the coordinator to encourage the participant to continue on study.

The protocol also stated that blood samples were taken at the time of grade 3-4 toxicities or outcome events "to assess protocol compliance."

According to the ERSMAC report, 29.2% of women on placebo and 33.8% of women on tamoxifen withdrew from the study.

Table 10. Reasons for being off therapy in violation of protocol procedures: Entire study population (Table 6, ERSMAC report)

Reason	Placebo		Tamoxifen	
	No. Pts.	%	No. Pts.	%
Never started	35	0.5	38	0.6
CW/LTFU*	631	9.4	636	9.5
Started & stopped:	1293	19.3	1582	23.7
Non- medical	643	9.6	583	8.7
Medical	650	9.7	999	15.0
Total	1960	29.2	2256	33.8

*Consent withdrawal/lost to follow up

The percent of women who consented but never took study medication and who withdrew consent or were lost to follow up is similar between the two treatment arms.

The ERSMAC report states that there was no difference in the percent of women who went off study by age:

Table 11. Reasons for being off therapy in violation of protocol procedures: By age at randomization (Table 7, ERSMAC report)

Age group (years)	Reason	Placebo		Tamoxifen	
		No. Pts.	%	No. Pts.	%
< 50	[redacted]	N=2638		N=2623	
	Never started	15	0.6	12	0.5
	CW/LTFU	266	10.1	275	10.5
	Started & stopped:	507	19.2	638	24.3
	Non-medical	243	9.2	239	9.1
	Medical	264	10.0	399	15.2
	Total	788	29.9	925	35.3
50-59	[redacted]	N=2052		N=2059	
	Never started	12	0.6	10	0.5
	CW/LTFU	184	9.0	185	9.0
	Started & stopped:	400	19.5	481	23.4
	Non-medical	203	9.9	151	7.3
	Medical	197	9.6	330	16.0
	Total	596	29.0	676	32.8
≥ 60	[redacted]	N=2017		N=1999	
	Never started	9	0.4	16	0.8
	CW/LTFU	181	9.0	176	8.8
	Started & stopped:	386	19.1	463	23.2
	Non-medical	197	9.8	193	9.7
	Medical	189	9.4	270	13.5
	Total	576	28.6	655	32.8

Reasons for withdrawing (never started, withdrew consent/lost to follow up, started and stopped for medical or non-medical reasons) did not differ among the protocol-specified age groups.

For women who began study therapy and then stopped therapy, withdrawals are summarized as follows:

Table 12. Reasons for starting and discontinuing therapy (Table 8, ERSMAC report)

REASON FOR BEING OFF THERAPY	PLACEBO (N=1293; 19% OF STUDY POPULATION)		TAMOXIFEN (N=1582; 24% OF STUDY POPULATION)	
	No. Pts.	%	No. Pts.	%
Non-medical	**643**	**9.6**	**583**	**8.7**
Medical:	**650**	**9.7**	**999**	**15.0**
Hot flashes	101	1.5	208	3.1
Anxiety	52	0.8	81	1.2
Weight gain	25	0.4	28	0.4
Headache	32	0.5	18	0.3
Skin itching/rash	24	0.4	21	0.3
Vaginal dryness	23	0.3	30	0.4
Vaginal discharge	5	0.1	31	0.5
Menstrual irregularity	29	0.4	30	0.4
Nausea	20	0.3	15	0.2
Other	339	5.1	537	8.0

The increased frequency of hot flashes and vaginal discharge observed on the tamoxifen arm could be anticipated from prior clinical experience with the drug.

"Other reasons" for stopping therapy included:

Table 13. Other reasons for stopping therapy

Vomiting	Diarrhea
Constipation	Bladder control difficulty
Genital itching	Pain with intercourse
Cramps	Breast tenderness
Ringing in ears	General aches and pains
Joint pains	Chest pain
Swelling of hands or feet	Difficulty breathing
Weight loss	Decreased appetite
Forgetfulness	Difficulty concentrating
Dizziness/faintness	Numbness/tingling
Hemorrhage	Infection/sepsis
Stomatitis	Hematuria
Alopecia	Pulmonary
Cardiac dysrhythmia	Cardiac function
Hypertension	Hypotension
Allergy	Other
Neuro complaints*	

*Neurosensory, neuro-motor, neuro-cortical, neuro-cerebellar, neuro-hearing complaints

Overall, each event was uncommon, representing less than 0.1% of patients on an arm, with the exception of infection/sepsis and alopecia. Infection/sepsis occurred in 12 patients on placebo (0.2%) and in 19 participants on tamoxifen (0.3%). Alopecia occurred in 20 women on placebo (0.3%) and in 39 women on tamoxifen (0.6%). Other differences included hemorrhage, which occurred in 3 tamoxifen patients compared to none on placebo; swelling of the hands and feet was noted in 7 women on tamoxifen compared to 1 on placebo. Bladder control difficulty and genital itching each occurred in 5 women on tamoxifen and in 1 woman on placebo. Neuro-cortical and neuro-cerebellar problems each resulted in stopping therapy for 2 women on placebo and 5 women on tamoxifen. Allergy was reported in 1 woman on placebo and 4 on tamoxifen. Adverse events that were more commonly reported for placebo included dizziness, forgetfulness, diarrhea, and breast tenderness (9 vs. 5; 5 vs. 2; 8 vs. 1; 7 vs. 2 respectively).

The annual rates of non-compliance are summarized in the following table:

Table 14. Annual rates of non-compliance (Sponsor table 10, ERSMAC report, volume 3, page 21).

	Percent Non-compliance		
Year of Follow-up	**Planned level**	**Placebo**	**Tamoxifen**
1	16	10.7	13.1
2	14	11.5	13.1
3	10	7.7	8.4
4	10	5.5	6.5
5	10	3.5	5.0

Reviewer Comment:

1. The original submitted database did not permit an assessment of how many women interrupted therapy, why, and for how long. These data were submitted 7/28 and 8/3/98.

2. No information on pill counts or participant/staff assessment of compliance was provided. During a teleconference with Zeneca and the NSABP on July 23, 1998, this issue was discussed. The NSABP stated that it stopped collecting compliance forms because compliance was high. Additional information was supplied August 12, 1998. At the start of the study, a data form was used to elicit information directly from participants regarding compliance. The second section of this form, "Staff assessment of the participant's compliance", was based on an interview with the patient. Pill counts were also performed. In February 1996, the data from the compliance form were analyzed and found to correlate with the data obtained from pill counts. The form was then discontinued. A local institution substudy was designed to assess this issue. Preliminary findings are discussed in section 12.0.

3. The sponsor was asked whether blood samples were actually collected at the time of toxicities or events as per protocol, and whether they have been analyzed for tamoxifen blood levels. In a reply dated August 12, 1998, the NSABP reported that "considerably less" than 50% of the event blood samples were collected. Because of the high compliance level of the participants and the striking study results, this blood will be used for research purposes. Tamoxifen blood levels will not be measured.

4. Additional information on compliance was supplied in the draft manuscript of the trial.

- 137 women on each arm provided no follow-up information (2.0% and 2.1% of the women on placebo and tamoxifen respectively)
- 7.2% of the women on each arm withdrew consent, but were followed until this point, and
- 1.4% of women on each arm were lost to follow-up
  - These numbers do not match those reported in Table 10, but are close (8.6% cited in the manuscript compared to 9.5% in the ERSMAC report)

- 21.4% discontinued treatment for non-protocol-specified reasons (Tables 12 and 13 document these reasons for 19% and 24% of the participants in the two treatment arms)

5. Reviewer summary:

The non-compliance rates, as expected, were higher for the tamoxifen group than for the placebo group. The difference in non-compliance is due almost entirely to medical reasons for discontinuation. No difference in the reasons for non-compliance was observed among the 3 age groups. However, both groups had a higher level of compliance than that assumed in the protocol design.

Hot flashes, anxiety, and vaginal discharge were the medical symptoms that accounted for the difference in dropout between the arms.

6. The sponsor was asked to provide information on the number of women who completed 5 years of drug therapy.

Table 14a. BCPT participants receiving 5 years of therapy

	Placebo	*Tamoxifen*	*Total*
No. randomized on or before 2/1/93	*2879*	*2860*	*5739*
No. completed 5 years of therapy	*1782*	*1596*	*3378*

Overall, 27% of the women randomized to placebo and 24% of the women randomized to tamoxifen completed 5 years of therapy.

8.7.2 Treatment delays/windows

Per protocol, there were no dose reductions permitted. Dose delays were outlined in the protocol for hepatic toxicity, hematologic toxicity, and other toxicities.

Reviewer Comment:

1. No information about changes in dosing or dosing delays was available for review.

8.8 Endpoints

The study was designed to evaluate several efficacy and safety endpoints. At the occurrence of any event, 40 ml of blood was to be collected and stored centrally

8.8.1 Breast Outcome Measures

Breast outcomes and the required supporting documentation were defined as follows:

- Invasive breast cancer:
  - Pathologic diagnosis of invasive breast cancer
  - Slides/blocks of tumor tissue submitted to NSABP (fresh frozen tissue if possible)
- Other breast outcomes
  - Pathologic diagnosis of non-invasive breast cancer or atypical hyperplasia
  - Slides/blocks of tissue submitted to NSABP

In order to maximize collection of breast outcome events, a mammogram was required annually and submission of the report was mandatory (protocol section 9.4.2). All breast biopsy and cytology results, benign or malignant, were required to be recorded on an event form. Mammogram reports, operative reports, pathology reports, and slides were required for positive or suspicious findings.

Reviewer Comment:

1. During review of the CRFs, breast biopsy and cytology reports were generally included. Mammogram reports were found uncommonly unless a biopsy was performed; the reports were then submitted as supporting documentation.

2. It is therefore not possible to evaluate whether all events were recorded in the database.

8.8.2 Cardiovascular outcome measures

The first safety endpoint was the effect of tamoxifen compared to placebo on cardiovascular mortality. All cardiovascular events, including arteriosclerotic vascular disease, and mortality from any atherosclerotic event were reported on an event form; additional primary source documentation was to be forwarded to the NSABP. Participants aged 50 or older at the time of study entry were required to have a baseline ECG and repeat studies at 3 years and 6 years. Copies of the ECGs were to be forwarded to the NSABP. All patients were required at baseline to have evaluation of fasting cholesterol and triglyceride levels. A subgroup of patients had continued evaluation of lipid profiles. These data are not available for review.

Outcomes of interest included:

Cardiovascular death:

- Definite vascular death: Any death due to myocardial infarction, stroke, pulmonary emboli, or other vascular event (i.e., ruptured aorta). Includes sudden deaths without any other cause
- Presumed vascular death: Any deaths without any clear nonvascular cause

Fatal and nonfatal myocardial infarction:

- Q-wave MI: Must have 2 of 3 criteria:
  - Characteristic symptoms of chest pain
  - New significant Q-wave on standard 12-lead ECG

- Significant elevation of serum enzymes
  - Elevation of CPK-MB to twice the upper limit of normal within 36 hours of onset of acute symptoms of MI
  - Reversal of LDH1/LDH2 ratio within 5 days of onset of acute symptoms of MI
  - CPK total at least twice the upper limit of normal for the lab
  - SGOT, LDH, or other cardiac enzymes at least twice the upper limit of normal for the lab
- Non-Q-wave MI: New and persistent ST changes on the ECG with significant enzyme elevation in the presence or absence of chest pain
- Fatal MI: Death within 7 days of MI or sudden cardiac death

Secondary events:

Secondary vascular events required submission of the hospital discharge summary and death certificate (when applicable). The following definitions were used:

- Acute ischemic syndromes
  - Probable MI: Presence of a new Q-wave on ECG; chest pain or significant enzyme elevation may not be present
  - Unstable angina: Angina pectoris requiring hospitalization
- Severe angina: Characteristic chest pain upon exertion that requires either PTCA or CABG
- Stroke: Presence of neurologic deficits that persist > 24 hours.
  - Occlusive
  - Hemorrhagic
  - Fatal: Death within 7 days of a stroke
- Transient ischemic attacks (TIA):presence of neurologic deficits for < 24 hours
- Pulmonary embolism: Clinical diagnosis confirmed with a V-Q scan
  - Definite: Clinical diagnosis with positive V-Q scan
  - Probable: Clinical diagnosis with a suspicious V-Q scan
- Deep vein thrombosis: Clinical diagnosis confirmed with venography, venous Doppler, venous duplex imaging, or fibrinogen scan
  - Definite: Clinical diagnosis with positive study
  - Probable: Clinical diagnosis with suspicious study
  - Normal: Clinical diagnosis with negative study
- Peripheral vascular disease: Significant PVD requiring vascular surgery

Reviewer Comment:

1. ECGs were present in the CRFs only as documentation of an event. Baseline and routine follow-up studies were not part of the CRF.

2. Lipid values were not found in the CRF or in the electronic database. The NSABP reported in a teleconference 7/23/98 that values were obtained at baseline only.

8.8.3 Fractures

The second safety endpoint was the effect of tamoxifen compared to placebo on the incidence of fractures.

- Fractures, regardless of site or cause, were to be reported on an event form.
- Submission of the X-ray report and supporting documentation (hospital summary, operative report, etc) was required.

Although not specifically stated in this section, section 2.3 of the protocol (Justification for the study design) indicated that hip and wrist fractures were the primary fracture sites of interest, because of their high prevalence, their associated morbidity, and the ability of investigators to reliably document them as events. Information on vertebral fractures will be collected, but they will not be included as events because there is no agreed-upon definition of a vertebral fracture, many vertebral fractures are unknown to the patient, and methods for determining vertebral fractures are costly and/or are not reproducible.

8.8.4 Endometrial cancer

All participants who had not had a prior hysterectomy with bilateral salpingo-oophorectomy were required to have a yearly gynecologic examination. Coordinators were to ask participants about menstrual irregularities at each visit and refer them for gynecologic consultation if this symptom was present. Refusal to comply with recommended evaluation was a reason for discontinuation of protocol medication.

If a participant was diagnosed with endometrial cancer, the following information was required for submission:

- Event form, for cancer, hyperplasia, occurrence of an endometrial biopsy or cytology
- Operative report
- Pathology report
- Tumor slides or blocks

Reviewer Comment:

1. Most CRFs contained extensive documentation of serial gynecologic evaluations. This information was requested in electronic format. It was sent 7/29/98 and scheduled screenings were included and coded as negative, hyperplasia, or cancer. It is therefore not possible to assess how many procedures were performed on each arm of the trial for dysfunctional uterine bleeding or other gynecologic symptoms. It is also not possible to evaluate the incidence of other gynecologic events that have been associated with tamoxifen in the literature, such as endometrial polyps.

8.8.5 Cancers other than breast cancer

- All cancers other than breast cancer were to be reported on an event form
- A copy of the pathology report and slides of the tumor were to be submitted for review; fresh frozen tumor specimens were requested when possible.

Reviewer Comment:

1. Other cancers appeared to be extensively documented in the CRFs.

8.8.6 Mortality

Mortality from any cause was to be reported on the Report of Death form. A death certificate was to be submitted, as well as a hospital discharge summary if the patient was hospitalized at the time of death.

8.8.7 Visual changes

Information was collected by asking patients at each visit about visual changes and ophthalmologic events. Forms were submitted; no primary documentation was obtained.

Reviewer Comment:

1. Electronic data on cataracts, cataract surgery, and macular degeneration were submitted at FDA request 7.28/98 and 8/4/98.

8.8.8 Quality of life

Quality of life forms were completed at baseline and at each follow-up visit.

Reviewer Comment:

1. At FDA request, the quality of life analyses were submitted 7/31/98 in lieu of primary data (except for depression scores, which were provided). NSABP assured the Agency that extensive analyses had been performed and no significant differences were noted across treatment arms. For further details, see Section 11.0, Quality of Life.

8.9 Statistical considerations

There were several objectives in this trial. The hypotheses, in order of importance, for the statistical calculations were:

- to determine if tamoxifen decreases the incidence of invasive breast cancer
- to determine if tamoxifen decreases the mortality from invasive breast cancer
- to determine if tamoxifen decreases the mortality from coronary heart disease

- to determine if tamoxifen decreases the occurrence of fatal and non-fatal MI
- to determine if there is a difference in the rate of fracture between the placebo and tamoxifen groups

The original calculations yielded a sample size of 16,000 women. This sample size was estimated to provide 80% power to detect a 40% reduction in the incidence of breast cancer after 9 years, using a one-sided alpha. The following assumptions were made in this calculation:

- Use of a one-sided alpha
- 10% of participants per year randomized to receive tamoxifen would withdraw from treatment, and that 1% of patients per year would be lost to follow up
- Incident breast cancer rates were based on 1983-1987 SEER data adjusted by age-specific estimates of the increase in risk observed in women who participated in a breast cancer screening program; age and race distributions were based on 1986 U.S. census data
- Person-years were adjusted for competing causes of death using U.S. age-and race-specific death rates. In order to adjust for the possibility that the study population may have a lower death rate from some causes than does the general population, the U.S. death rate for coronary heart disease was decreased by 30% and all other non-breast cancer causes by 10%
- The targeted 40% reduction in breast cancer risk was based on literature reports of a 50% reduction in the risk of contralateral breast cancers by tamoxifen. After adjustment for a 10% per year withdrawal rate, the NSABP estimated that the 40% reduction in failure rate is approximately equal to a 31% average reduction in failure rate.

The power calculations for the other endpoints of the trial are listed below. The estimates are based on an accrual rate of 8000 participants per year for 2 years and a one-sided alpha for all outcomes except fractures. For fractures, an alpha of .01 in a two-sided test was used.

Table 15. Minimum detectable % reduction in hazard for specified power and follow-up periods (sponsor table 12, from the protocol, section 15.1)

	5-YEAR FOLLOW-UP		8-YEAR FOLLOW-UP	
STATISTICAL POWER	.8	.9	.8	.9
ENDPOINT				
Breast Cancer Incidence	35	40	29	33
Breast Cancer Mortality	68	76	56	62
Coronary Deaths	50	56	38	43
Fatal/Nonfatal MI	34	38	26	29
Hip and Colles' Fractures	30	33	23	27
Hip Fracture	51	57	40	44

The risk ratios used for these estimates were derived from those observed in the BCDDP trial, applied to U.S. breast cancer mortality rates. In order to adjust for the fact that the selection criteria for breast cancer risk may have resulted in a lower than average risk for cardiovascular events, the cardiovascular mortality rate was reduced by 30% below the age-race-sex specific rates for the US population, and the incidence of myocardial infarction was lowered 20% below the Framingham incidence. Age-specific fracture rates for hip and Colles' fractures were derived from the literature.

A recalculation of the sample size was planned after accrual of 25% of the participants, in order to ensure that the risk of the enrolled women matched the assumptions. Provisions were made for amending the protocol in terms of sample size if necessary. The sample size was amended September 30, 1996 to 13,000 women. The amendment was due to the following observations: the placebo group had an annual breast cancer incidence of 0.00668, twice the anticipated rate, suggesting that women in this study were at higher risk than originally estimated; the average annual lost-to-follow-up rate was 3%, 3 times higher than expected; actual non-compliance rates were 16% for the first year of follow-up and 14% for the second year, with an estimated non-compliance rate of 10% for years 3-5; and the need to consider a linear depletion from 40% to 20% in the reduction of breast cancer incidence in years 6-10, after stopping tamoxifen. The original calculations did not take into account a possible loss of treatment effect. When these observations were incorporated, the sample size of 13,000 was determined to provide 82% power to detect a 40% reduction in the occurrence of breast cancer.

Accrual was stratified by age (35-49, 50-59, 60+), race (black, white, other), history of LCIS (yes/no), and breast cancer relative risk (< 2.5, 2.5-3.9, >4.0). An adaptive randomization scheme, using the biased coin method of Efron, was used to avoid extreme inequalities in treatment assignment within a clinical center.

Interim analyses were planned, for both safety and efficacy. Interim safety reviews were planned every 6 months. Blinded data on accrual, compliance, distribution of baseline risk factors, toxicity, and data quality were provided to the ERSMAC for review. Specific toxicity parameters included hot flashes, nausea, and vomiting. Information was provided for the entire study population as well as by center. The occurrence of endometrial cancer and thromboembolic events was evaluated using a rapid reporting system.

Interim analyses for the primary endpoint of interest, the occurrence of invasive breast cancer, were performed yearly using Fleming stopping rules. A maximum of 8 interim analyses was planned with bounds that used less than 5% of the critical region. The test statistic used was the two sample t-test applied to total incidences cases, assuming a Poisson distribution.

Analysis for other endpoints of interest were planned yearly, but would not begin until 50 breast cancer cases were identified. If the analyses met Fleming stopping rules, the life-table curves and analyses were planned per protocol for presentation to the ERSMAC in order to determine whether the study should be stopped early.

If the observed incidence of breast cancer in the tamoxifen group was greater than that observed in the placebo group, a standardized difference of 1.4 was planned as the screening point that would prompt a review by the ERSMAC.

Analysis of cardiovascular mortality was performed yearly with the same boundaries specified for breast cancer mortality monitoring. At any interim analysis, where the test statistic used to compare coronary heart disease mortality in the two groups exceeds the boundary values, the results were to be reviewed by the ERSMAC. The statistical power for cardiovascular endpoints was recognized as dependent on the age of the study population. If the number of postmenopausal women was less than 10,000 at the end of the trial, adjustment in accrual patterns would be considered.

An intent-to-treat analysis was planned for all endpoints (randomized, not actual, treatment), based on the assumption of a Poisson distribution, using a linear model for comparison between tamoxifen and placebo groups. The Poisson assumption was made for breast cancer mortality, coronary heart disease mortality, MI, and fractures. Kaplan-Meier life tables were planned as a secondary analysis for the patterns of events and would be compared using logrank tests and the Cox proportional hazard model. The effect of noncompliance and withdrawal on the estimated treatment effects was planned for evaluation. An attempt at modeling a dose-response relationship would be made.

In a facsimile dated June 5, 1998, Joseph Constantino further outlined the statistical methods used. Average annual event rates for the study endpoints were calculated by treatment group and were defined as the number of observed events divided by the number of observed event-specific person years of follow up. Fisher's exact test was used to determine p values for the rates of invasive and non-invasive breast cancer. Relative risk ratios were also calculated for events. The confidence limits for the relative risk were based on the variance for the log of the relative risk and were calculated using the exact method. The method of Korn and Dorey was used to account for competing risk due to death in the analysis of cumulative incidence rates by follow up time.

Reviewer Comment:

1. The study was powered to detect a benefit of tamoxifen, but was not powered to detect a detrimental effect of treatment with tamoxifen.
2. The sample size re-calculation was planned to occur after 25% of the patients had been accrued. Based on the original target of 16,000 women, the sample size evaluation should have been performed after accrual of 4000 women, midway through the first year (accrual from 6/92 through 5/93 was 7154, according to the amendment). The amendment was not made until 9/30/96. The delay was due to the need for prolonged follow-up of the participants to fully evaluate risk levels.
3. The number of women on study aged 50 or older, evaluated by an Access query of the database, was 8127, 4069 on placebo and 4058 on tamoxifen, rather than the 10,000 calculated as the approximate sample size to evaluate cardiovascular events. The failure of tamoxifen to reduce myocardial infarction may be due to power considerations rather than a true lack of effect.

   The sponsor was asked to comment on this observation by the medical reviewer. The sponsor indicated that power might be an issue (August 19, 1998). However, the number of events in each treatment group was comparable for the 4 heart disease endpoints. While lack of power may have contributed to the failure to detect an effect of tamoxifen, the support for such

an effect would have been stronger if there was at least some trend of fewer events in the tamoxifen arm.

Tony Koutsoukos, Ph.D., the Statistical reviewer, performed a power calculation based on the accrued sample size for women aged 50 or older. From Table 8 of the protocol, page 2.19, the five year average (ages 50-70) of the projected probabilities of coronary heart disease is 2.12%. This rate falls between the rates of coronary artery disease reported for women age 60 and age 65. An assumption of a 35% reduction of this coronary heart disease rate for the treatment group was made, based on the statements in the protocol used in the original sample size/power calculations. With a two-sided level of significance of 5% and about 4060 patients in each treatment group, there will be approximately 73% power to detect a 35% reduction of this coronary heart disease rate (from 2.12% to 1.38%) for the treatment group.
(If n=4000 per group, then power=72%)
(If 30% reduction, then power=52%)

These calculations suggest that while an effect of tamoxifen cannot be excluded, the accrued sample size should have been large enough to identify a trend to reduction in cardiovascular events.

4. We have not seen analyses that attempt to model a dose-response relationship. Because no dose reductions were permitted, we assume this statement refers to duration of therapy rather than a dose question.
The sponsor replied 8/19/98 that in such an analysis for P-1, dose would be measured as time on drug. Such an analysis is biased, because there are usually differences in outcome between those who adhere to therapy and those who do not (even on a placebo arm). This analysis was intended only if no treatment effect was found for the invasive breast cancer endpoint. Because a substantial treatment effect was observed, no dose-response analysis is planned.
5. All randomized participants with follow-up were included in the analyses, according to a fax sent by Joseph Costantino 6/5/98. The NSABP analyzed 13,114 women (6570 on placebo and 6544 on tamoxifen). The FDA generally analyzes data in a true intent-to-treat analysis, where all randomized participants are included, regardless of whether or not they received drug or returned for follow-up. Given the size of the study, it is unlikely that this difference in the definition of "intent-to-treat" will alter the results. All women who received drug should be evaluable for toxicity, regardless of length of follow up.

9.0 Efficacy Review

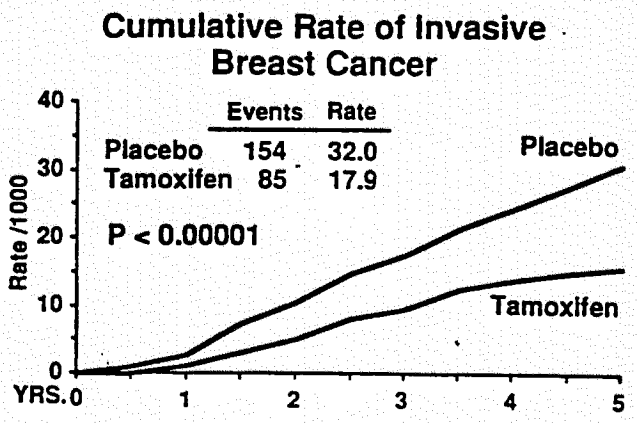
9.1 Invasive breast cancers

Two hundred thirty-nine invasive breast cancers occurred, 154 on placebo and 85 on the tamoxifen arm. At the interim analysis performed on events through January 31,

1998, the p-value calculated with the O'Brien-Fleming stopping rule was p=0.00017. The difference between treatment arms was significant with a p-value of 0.0000055.

Overall, there is a 45% reduction in the relative risk of developing invasive breast cancer. In the technical report, it is stated that the benefit was apparent early in the course of treatment and is still observed through all periods of follow-up. A copy of a slide used by the NSABP at ASCO and at the June NSABP meeting illustrates this point:

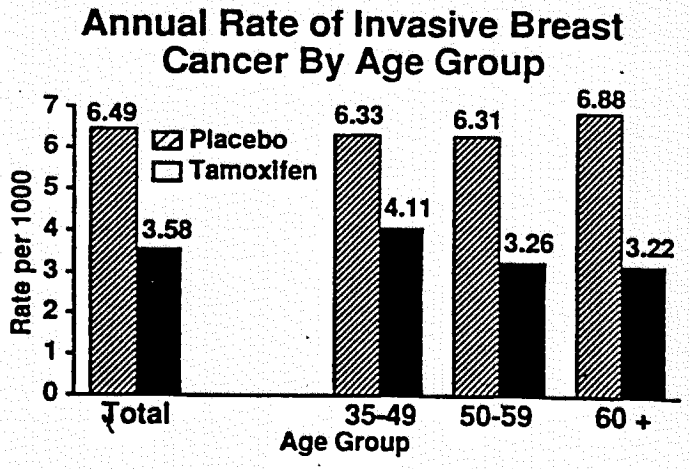
Figure 1.



There is no survival difference in disease-specific mortality: there have been 5 breast cancer deaths in the placebo group and 3 in the tamoxifen-treated group.

According to an NSABP presentation at ASCO/NSABP meeting, tamoxifen was effective in preventing breast cancer in all age groups:

Figure 2.



The overall reduction in annual rate of cancer was 45%, based on this slide. For women aged 35-49, the reduction was 35%; for women aged 50-59, the reduction was 48%; and for women aged 60 or greater, the reduction was 53%.

The NSABP technical report presented the data used in the above bar graph in another form, with 95% CI: