International Breast Cancer Study Group Statistical Center

TEXT (Trial 25-02/BIG 3-02) and SOFT (Trial 24-02/BIG 2-02)

Statistical Analysis Plan

2017 Update

Meredith Regan, ScD

May 2017	Complete for efficacy, AE, protocol treatment red text reflects changes for the update analyses since the first analyses blue text reflects edits/notes since the analysis started
Nov 2017	For non-protocol therapy notes (noted in blue)

1. Introduction

1.1 Background

In 2003 IBCSG initiated a suite of three complementary tailored treatment investigations, the SOFT, TEXT and PERCHE trials, designed to answer questions concerning adjuvant treatment for premenopausal women with endocrine-responsive early breast cancer: 1) What is the role of ovarian function suppression (OFS) for women who remain premenopausal and are treated with tamoxifen? 2) What is the role of aromatase inhibitors for women treated with OFS? 3) What is the role of chemotherapy for women treated with OFS plus oral endocrine therapy?

The conduct of these randomized phase III trials required world-wide participation through collaboration of the Breast International Group (BIG) network and the North American Breast Cancer Groups. Pfizer is the primary pharmaceutical industry partner and Ipsen provides triptorelin outside of North America.

Over a 7.5-year period from 2003 to 2011, 5742 premenopausal women were enrolled at over 500 centers in 27 countries on 6 continents in:

TEXT (Tamoxifen and Exemestane Trial): designed to determine the role of AIs for women who receive OFS from the start of adjuvant therapy;

SOFT (Suppression of Ovarian Function Trial): designed to determine the role of OFS and the role of AIs for women who remain premenopausal after completion of (neo)adjuvant chemotherapy, or who are premenopausal following surgery and tamoxifen alone is a reasonable treatment option;

PERCHE (Premenopausal Endocrine-Responsive Chemotherapy): designed to determine the value of adding chemotherapy to combined endocrine therapy with OFS plus oral endocrine therapy.

TEXT and SOFT successfully enrolled the targeted number of patients. PERCHE closed prematurely in 2006 with only 29 patients enrolled (25 of 29 were co-enrolled in TEXT; (Regan et al., Ann Onc 2008).

Completion of TEXT and SOFT enrollment was anticipated within 5 years and first reporting about 7 years after the trials' initiation. However the characteristics of the enrolled patients differ from those anticipated in the protocols, and patients' outcomes are better than expected, necessitating an adaptation of the trials' analysis plans, in protocol amendments released in 2011. The original designs of TEXT and SOFT and the adaptations to overcome these challenges and ensure timely answers to questions concerning adjuvant treatment for premenopausal women with endocrine-responsive early breast cancer were described recently (Regan et al., Breast 2013).

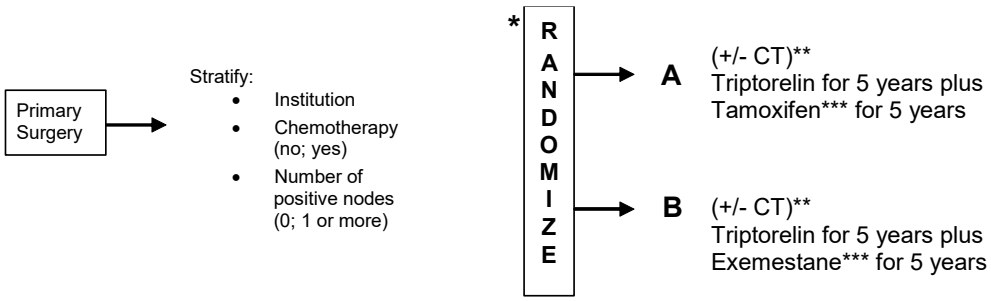
1.2 Trial Designs

The designs of TEXT and SOFT are summarized below.

TEXT (IBCSG 25-02 / BIG 3-02)

Title:	**Tamoxifen and Exemestane Trial (TEXT):** A phase III trial evaluating the role of exemestane plus GnRH analogue as adjuvant therapy for premenopausal women with endocrine responsive breast cancer.
Patient Population:	Premenopausal women (estradiol (E2) levels in the premenopausal range) with histologically proven, resected breast cancer with ER and/or PgR positive tumors (ER and/or PgR ≥ 10%).
Entry:	Patients should be randomized within 12 weeks after surgery prior to commencing any adjuvant systemic therapy.
Activation Date:	04Aug03 (First patient randomized 7Nov03)
Target Accrual:	2639 patients
Closure Date:	11Mar11 (Last patient randomized 7Apr13)
Final Accrual:	2672 patients

TEXT Schema:



* Randomization prior to receiving any adjuvant systemic therapy

**CT (chemotherapy) is determined by investigator/patient choice and if used, should begin at the same time as triptorelin.

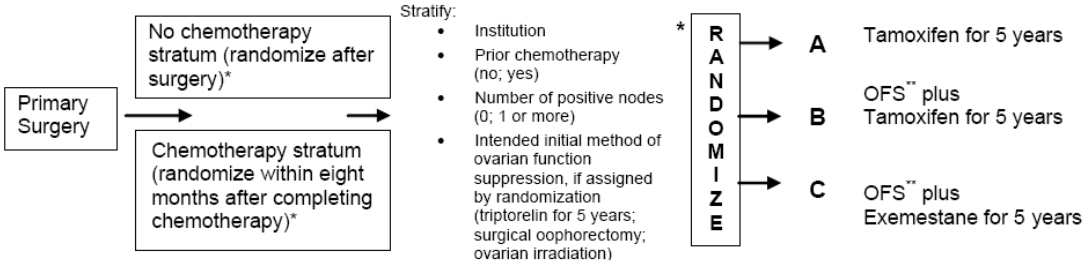
***Tamoxifen or exemestane should start after adjuvant chemotherapy has been completed or approximately six to eight weeks after the initiation of triptorelin, whichever is later.

IBCSG 25-02 (TEXT) is an international, non-blinded, randomized phase III trial designed to investigate the efficacy the aromatase inhibitor (AI) exemestane with OFS, achieved by long-term use of GnRH analogue, compared with tamoxifen+OFS. TEXT focuses the AI question on premenopausal patients whom the physician feels OFS is most appropriate from the start of adjuvant therapy. Eligibility required enrollment within 12 weeks of definitive surgery and excluded patients who had already received any (neo)adjuvant chemotherapy or endocrine therapy. Randomization used 1:1 allocation and stratified by whether adjuvant chemotherapy was planned (no; yes), and number of positive nodes (0; 1 or more). Patients would either receive no chemotherapy or commence chemotherapy at the same time as GnRH analogue is initiated. Patients had to use GnRH analogue for at least 6 months and then could change to permanent OFS with surgery or radiation at any time thereafter.

SOFT (IBCSG 24-02 / BIG 2-02)

Title:	**Suppression of Ovarian Function Trial (SOFT):** A phase III trial evaluating the role of OFS and the role of exemestane as adjuvant therapies for premenopausal women with endocrine-responsive breast cancer.
Patient Population:	Premenopausal women (estradiol (E2) levels in the premenopausal range†) with histologically proven, resected breast cancer with ER and/or PgR positive tumors (ER and/or PgR ≥ 10%) who have received either no chemotherapy or remain premenopausal following completion of adjuvant and/or neoadjuvant chemotherapy.
Entry:	Patients who do not receive chemotherapy should be randomized within 12 weeks after surgery; such patients must have E2 levels in the premenopausal range following surgery. Patients who have received adjuvant and/or neoadjuvant chemotherapy should be randomized within 8 months of the final dose of chemotherapy as soon as premenopausal status is confirmed; such patients must have E2 levels in the premenopausal range between 2 weeks and 8 months after the final dose of chemotherapy.
Activation Date:	04Aug03 (First patient randomized 17Dec03)
Target Accrual:	3000 patients
Closure Date:	31Jan11 (Last patient randomized 27Jan11)
Final Accrual:	3066 patients

SOFT Schema:



* Patients may have received tamoxifen or an anti-aromatase agent prior to randomization
** OFS = ovarian function suppression (triptorelin for 5 years OR surgical oophorectomy OR ovarian irradiation)

IBCSG 24-02 (SOFT) is an international, three-arm, non-blinded, randomized phase III trial designed to investigate the role of OFS and the role of the AI exemestane, originally planned with three primary comparisons: tamoxifen+OFS versus tamoxifen alone; exemestane+OFS versus tamoxifen alone; and exemestane+OFS versus tamoxifen+OFS. SOFT focuses the OFS question on those who biologically would be most likely to benefit, i.e., women with endocrine-responsive breast cancer with premenopausal status either after completion of (neo)adjuvant chemotherapy or following surgery alone.

Eligibility required enrollment either: (a) within 8 months of the final dose of chemotherapy once premenopausal status was confirmed by estradiol levels (e.g., patients with temporary chemotherapy-induced amenorrhea who regained premenopausal status within 8 months were eligible); or (b) within 12 weeks of definitive surgery if no adjuvant chemotherapy was to be given. Patients could have received adjuvant oral endocrine therapy (but not GnRH analogues) for up to 8 months prior to randomization. The 8-month criterion was an early protocol amendment (from 6 months) to overcome logistical challenges of enrolling a patient who presented after regaining premenopausal status at a 6-months post-chemotherapy standard-of-care visit.

Randomization used 1:1:1 allocation, and was stratified according to prior (neo)adjuvant chemotherapy (yes;no), lymph node status (0; 1+) and intended method of OFS (GnRH analogue, oophorectomy; ovarian irradiation). For patients randomized to receive OFS, the use of GnRH analogue, bilateral oophorectomy or bilateral ovarian irradiation was by patient preference and patients who began with GnRH analogue could opt to undergo surgery or irradiation at any time.

1.2.1 Design Commonalities

As a planned suite of trials, a majority of trial design features were common to both trials. Briefly, the trials enrolled premenopausal women with histologically-proven, resected, hormone receptor-positive (defined as ER≥10% and/or PgR≥10%) early invasive breast cancer. Premenopausal status was defined by estradiol levels in the premenopausal range according to institutional parameters. The tumor was to be confined to the breast and axillary lymph nodes without detected metastases elsewhere. Patients must have had proper local-regional treatment for primary breast cancer with no known residual loco-regional disease. Study visits were every 3 months during year one, every 6 months during the next 5 years, and yearly follow-up thereafter.

The oral endocrine therapy was either tamoxifen or the steroidal AI exemestane. OFS was by GnRH analogue triptorelin administered 4-weekly for 5 years, bilateral surgical oophorectomy, or bilateral ovarian irradiation (with biochemical confirmation of cessation of ovarian function after 2 months).

1.2.2 Original statistical design assumptions and sample size considerations

TEXT planned enrollment was 1845 patients. The design projected that 4.5 years of uniform accrual, plus 2.4 years of additional follow-up, would be sufficient to observe the target of 396 DFS events, which would provide 80% power to detect 25% reduction in hazard with exemestane+OFS versus tamoxifen+OFS (HR=0.75; 79.8% versus 74.1% 5-year DFS, respectively) using a 2-sided 0.05 α-level logrank test and assuming exponential distribution of DFS. Four interim analyses prior to the final analysis were planned. By November 2007, 2039 of the planned 1845 patients had enrolled, and enrollment was suspended. Because of the faster-than-expected enrollment rate and lower-risk characteristics of enrolled patients than anticipated, Amendment 2 (Jul08) re-opened enrollment with an increased target sample size of 2639 patients. A revised estimate of 80% 5-year DFS in the tamoxifen+OFS control group (with corresponding 25% reduction in hazard to 84.6% 5-year DFS for exemestane+OFS) was hypothesized based on the 2007 overview analysis of GnRH analogues in which the 5-year breast cancer recurrence was around 18% among patients treated with GnRH analogue plus tamoxifen. With the observed enrollment pattern and revised hazard rates, the increased sample size was projected to reach the target of 396 DFS events within 0.5 years of the original design, or 7.4 years since first enrollment.

SOFT planned enrollment was 3000 patients for the 3 arms. The design projected that 5 years of uniform accrual, plus 1.9 years of additional follow-up would be sufficient to observe the target of 783 DFS events (522 per pairwise comparison) to have 80% power to detect a 25% reduction in hazard relative to control 5-year DFS of 67% (HR=0.75; 74.1% versus 67.0% 5-year DFS; 2-sided α=0.0167). If tamoxifen+OFS would result in a 25% reduction in hazard to 74.1% 5-year DFS, then power was 68% to detect a further 25% reduction with exemestane+OFS to 79.8% 5-year DFS. Four interim analyses prior to the final analysis were planned.

From the outset, the protocols planned to combine the data of TEXT with the two arms of SOFT comparing exemestane+OFS versus tamoxifen+OFS. Differences in the two trials with respect to selection and treatment for women who received chemotherapy (i.e., TEXT enrolled patients following surgery and used concurrent GnRH analogue and chemotherapy, while SOFT enrolled patients who remained premenopausal following chemotherapy and initiated OFS after completion of chemotherapy) were taken into account in the combined analysis plan. The statistical power of such a combined comparison (at the two-sided α=0.05 level) would be at least 88%, 98% and 99% to detect a 20%, 25%, and 30% reduction in hazard, respectively, with exemestane+OFS versus tamoxifen+OFS under the protocol assumptions about accrual duration and additional follow-up.

1.2.3 Adaptations in the statistical design and analysis plans

As of October 2010, the overall DFS event rates—blinded to treatment assignment—were substantially lower than originally anticipated: approximately 1.7% and 2% per year versus the protocol-specified 6% and 8% per year in TEXT and SOFT, respectively. IBCSG projected an additional 7 and 13 years of follow-up to observe the targeted 396 and 783 DFS events in TEXT and SOFT, respectively (at median follow-up of 10.5 and 15 years). Increasing the sample size could hasten reaching the required events, but finances constrained this possibility.

The Steering Committee considered this delay to be unacceptably long (reporting 14 and 20 years after first enrollment versus 6.9 years originally-anticipated). The Committee decided to change the timing of analysis from “event-driven” to “time-driven” with a planned data cut-off during the third quarter of 2013, when the median follow-up should be at least 6 and 5 years in TEXT and SOFT, respectively. It was recognized that an analysis with fewer events than targeted would substantially reduce statistical power for the original protocol-planned primary objectives (approximately 60% in TEXT and 35% in SOFT to detect 25% reductions in hazards, assuming the October 2010 event rates continued). Therefore amendments of the TEXT and SOFT protocols (July 2011) revised the analysis plans for the first reporting of the trial objectives:

1. **AI Question**: the primary analysis comparing exemestane+OFS versus tamoxifen+OFS would implement the originally-planned combined analysis of TEXT and SOFT. The power of such a combined comparison (two-sided α=0.05 level) would be at least 95%, 84% and 63% to detect a 30%, 25% and 20% reduction in hazard, respectively, with exemestane+OFS.
2. **OFS Question**: the primary analysis from SOFT would focus on the unique comparison of tamoxifen+OFS versus tamoxifen alone, tested at the two-sided α=0.05 level. IBCSG estimated power to be at least 80%, 69%, 52% and 34% to detect 33.5%, 30%, 25% and 20% reductions in hazard, respectively, with tamoxifen+OFS.

These power calculations assumed a data cut-off in the third quarter of 2013 and persistence of the October 2010 DFS event rates, which project 250 DFS events in TEXT and 280 DFS events in SOFT (about 93 per group under the null hypothesis) at the time of data cut-off. The revised analysis plans removed planned interim efficacy analyses. The revised analysis plans removed planned interim efficacy analyses.

The Steering Committee's decision was endorsed by the IBCSG DSMC. These committees did not receive, nor did the IBCSG Statistical Center have knowledge of, outcome data according to treatment group prior to this decision. The first report of the combined analysis of the **AI Question** is anticipated in mid-2014. The report of the **OFS Question** from SOFT is anticipated in late 2014 after about 6 additional months of follow-up and a median follow-up of at least 5 years is reached.

Patient follow-up will continue and updates of efficacy results are planned approximately every two years after the first reports.

1.3 First Results

1.3.1 AI Question

After a median follow-up of 68 months (5.7 years), DFS events were reported for 514 (11%) of 4690 patients in the ITT population. Patients assigned E+OFS had statistically significantly reduced hazard of a DFS event vs. patients assigned T+OFS (HR=0.717; 95% CI, 0.602 to 0.855; P=0.0002). The estimated 5-year DFS was improved at 91.1% vs. 87.3% among patients assigned E+OFS vs. T+OFS. The reduction in hazard with E+OFS vs. T+OFS was consistently observed for secondary endpoints of breast cancer-free interval and distant recurrence-free intervals, although at this point in follow-up a reduction in hazard of death was not observed.

In this analysis, 30.2% of patients continue on protocol-assigned treatment. 30.6% and 29.4% of patients assigned to E+OFS and T+OFS have had grade 3-4 targeted AEs reported, of which hot flashes/flushes, musculoskeletal symptoms and hypertension were the most frequently reported

1.3.2 OFS Question

After a median follow-up of 67 months (5.6 years), DFS events were reported for 299 (14.7%) of 2033 patients in the ITT population. T+OFS non-significantly reduced hazard of a DFS event vs. T (HR=0.83; 95% CI, 0.66 to 1.04; P=0.10). The estimated 5-year DFS among patients assigned T+OFS was 86.6% (95% CI, 84.2% to 88.7%) versus 84.7% (95% CI, 82.2% to 86.9%) among patients assigned T. The relative effects of T+OFS vs. T were consistent between the two chemotherapy strata; as expected, the absolute difference at 5 years was more notable among patients who received prior chemotherapy. The reduction in hazard with T+OFS vs. T was consistently observed for the secondary endpoints of breast-cancer-free interval and distant recurrence-free interval. 106 (5.2%) deaths were reported, and on average there was a non-significantly reduced hazard of death; with all except 10 deaths among patients who received prior chemotherapy, the reduction in hazard was more apparent among this cohort of patients

At the time of the analysis, 26% of patients continued on protocol-assigned treatment. 31% and 24% of patients assigned to T+OFS and T have had grade 3-5 targeted AEs reported, of which hot flashes/flushes, musculoskeletal symptoms and hypertension were the most frequently reported.

1.3.3 Numbers of DFS events, deaths and median follow-up by trial

For the AI Question analysis, the TEXT ITT population had 285 DFS events and 101 deaths reported, after median follow-up of 6 years (data cut-off, 31Aug13).

At the time of the OFS Question analysis, the 3 arms of the SOFT ITT populations had 412 DFS events and 164 deaths reported, after median follow-up of 5.6 years (data cut-off, 31Mar14).

2. Efficacy Analysis Plans

2.1 Protocol Objectives

Based on revised statistical analysis plans in the 2011 amendments (i.e., SOFT Amendment 2 (24Aug11); and TEXT Amendment 3 (24Aug11)):

2.1.1 Primary Objective

AI Question: To evaluate the efficacy of exemestane+OFS compared with tamoxifen+OFS, across both trials combined, as adjuvant endocrine therapy for premenopausal women with endocrine-responsive early invasive breast cancer.

OFS Question: To evaluate the role of OFS+tamoxifen compared with tamoxifen alone, as adjuvant endocrine therapy for premenopausal women with endocrine-responsive early invasive breast cancer.

2.1.2 Secondary Objectives

- Assess overall survival, breast cancer-free interval, and distant recurrence-free interval
- Assess quality of life (see separate analysis plan)
- Investigate sites of first treatment failure
- Assess late side effects of early menopause
- Assess the incidence of second (non-breast) malignancies and causes of death without cancer event
- TEXT: Investigate patient and tumor features that may contribute to inter-individual variability of responsiveness to GnRH analogue plus exemestane and GnRH analogue plus tamoxifen (see separate analysis plan)

2.2 Updated Analyses

Rather than presenting results by trial, the analyses will continue to be organized as:

- AI Question (E+OFS vs T+OFS), jointly for the two trials;
  - Summarize by 4 cohorts
  - Summarize by trial also
- OFS Question (T+OFS vs T), reporting all three arms

The analysis will secondarily focus on the subgroup of HER2-negative disease. The main reason is the clinical relevance of separating the subgroups of HER2-negative and positive disease in current practice and trial design. We also previously observed heterogeneity of treatment effects according to HER2 status, and the HER2-postiive subgroup presents challenges because HER2-directed therapy changed during the conduct of the trials and HER2-directed therapy had different timing in TEXT and

SOFT chemotherapy cohorts (starting after randomization in TEXT, starting and mostly completing prior to randomization in SOFT).

2.2 Analysis Populations

The updated analysis will continue to use an intention-to-treat (ITT) approach. During the first analyses, there was one SOFT T+OFS patient who was included in the AI Question ITT population, but then excluded from the OFS Question ITT population (243007). In updated analyses, this patient will remain in the ITT populations as previously analyzed.

AI Question: The AI Question ITT population will include

- All patients randomized in TEXT (exemestane+OFS and tamoxifen+OFS);
- SOFT patients randomized to the two OFS-containing arms (exemestane+OFS and tamoxifen+OFS).

OFS Question: The OFS Question ITT population will include

- SOFT patients randomized to the two tamoxifen-containing arms (tamoxifen+OFS and tamoxifen alone).
- SOFT patients randomized to exemestane+OFS.

2.3 Endpoint Definitions

Endpoint definitions remain the same as in the first report.

Per IBCSG standard, the date of DFS event is the date a proven recurrence was first suspected. Since the first trial report, the 2425-RC Form SOP was clarified that in cases with multiple proven sites of first recurrence (within 2 months of each other), that the date first suspected should be the earliest of suspected dates from among those proven sites (which may not necessarily correspond to the site with the earliest proven date).

2.4 Follow-up

As in the first analysis, median follow-up is calculated from the Kaplan-Meier estimate of overall survival, with the event/censoring indicator inverted (i.e. alive as event and dead as censored).

In the first analysis, the visit cut-offs were 31Aug13 and 31Mar14. Updated results were anticipated to be presented every two years thereafter until the last patient enrolled has been followed for at least 10 years.

For the first update, concurrent visit cut-offs (of 31Dec16) and database locks (04May17) are planned. The visit cut-off is selected because almost all patients will have at least 6 years follow-up (except those final patients randomized during Q1 2011).

2.5 Tests and Estimates

As in the first analysis, the primary objectives will be investigated by comparing DFS between two treatment groups using two-sided stratified logrank test (H0: DFS1=DFS2; Ha: DFS1≠DFS2), with an overall experiment-wise α=0.05 (i.e., α=0.05 for AI Question analysis; and α=0.05 for OFS Question analysis). The test statistic and p-value will be taken from the stratified Cox PH model score test. Hazard ratios (AI Question: E+OFS / T+OFS; OFS Question: OFS+T / T) will be estimated from a stratified Cox PH model, with two-sided Wald 95% CIs. Kaplan-Meier estimates of the DFS distributions will be calculated for each of the treatment arms, with reporting of the 8yr DFS (MFU is 7.9y in SOFT and 9.1y in TEXT so 8yr estimates appropriate for both trials); the SEs will use Greenwood's formula and the pointwise 95% CIs will be obtained using complementary log-log transformation of the SDF.

The E+OFS vs T comparison will be included with estimation of HR (95% CI) for E+OFS / T, but no hypothesis testing is planned. The 8yr DFS would also be summarized.

In the first OFS Question analysis, the primary analysis model included only the 2 treatment groups; it is now planned that the model for HR estimates will include all 3 treatment groups, using contrasts to obtain the pairwise score tests.

We will check the proportional hazards assumption by visually assessing the plot of log(-log(survival)) versus log of survival time for parallelism. This will be done overall, and according to strata.

2.5.1 Stratification variables for logrank tests and HRs

AI Question analysis:

- Trial (TEXT; SOFT);
- Prior/Intended chemotherapy (no; yes);
- Number of positive nodes (0; 1+).

OFS Question analysis:

- Prior chemotherapy (no;yes);
- Number of positive nodes (0; 1+), and

The analyses will use the originally-defined strata variables, even if the actual value on the A form has been changed within the database since the first analysis (any such changes will be noted).

Since the first trial report, it was determined that the TEXT chemo strata hadn't paid adequate attention to value of patients enrolled in PERCHE whose chemo receipt was by random assignment. Upon review, those randomized to yes-chemo were all 12/12 in the yes-chemo strata, though only 11/13 of those randomized to no-chemo were in the no-chemo strata. As above, we will use the originally-defined strata variables but the observed adj_ct variable now says "no chemo" for all 13 pts.

Note about the stratification of OFS Question analysis: the two methods of ovarian ablation, oophorectomy and ovarian irradiation, were planned to be grouped as the third stratification factor because few patients had specified irradiation. Even grouped, having the 3 stratification factors (for 8 strata total) resulted in some small strata sizes with sparse events especially once considering the chemotherapy cohorts separately or other subgroups and once looking at outcomes with fewer events. As such, the outcome analyses will NOT include the third stratification factor, but the endpoint analyses in the overall ITT population will be repeated with the third stratification factor included to understand ramifications of this decision.

2.5.2 Cohorts

The TEXT and SOFT trial populations include four cohorts of patients:

- Patients in TEXT who were randomized within 12 weeks of surgery and for whom chemotherapy was not planned;
- Patients in SOFT who did not receive chemotherapy and were randomized within 12 weeks of surgery;
- Patients in TEXT who were randomized within 12 weeks of surgery and for whom chemotherapy was planned to begin after randomization, concurrently with GnRH analogue;
- Patients in SOFT who are premenopausal after prior (neo)adjuvant chemotherapy and are randomized within 8 months of completion chemotherapy.

Therefore in addition to having trial and chemotherapy as stratification factors, it is of interest to estimate the treatment effect within cohorts of the trials' patient populations. Note, the two cohorts of TEXT and SOFT who will not receive chemotherapy may differ, considering that in TEXT all patients would receive OFS whereas in SOFT the patients may have been randomized not to receive OFS.

(n.b., We will use the chemotherapy stratification factor, which in TEXT is intended chemotherapy and therefore may be different from actual chemotherapy receipt, in a few patients).

AI Question: In addition to the overall treatment comparison, the HRs (95% CI) will be estimated within each of the 4 trial-by-chemotherapy cohorts separately. The 4 cohort HR estimates will be obtained from adding the cohort-by-treatment interaction and using contrasts to estimate HRs and CIs within each of the 4 cohorts.

The HRs (95% CI) will also be estimated by trial. To obtain these estimates, we will refit the primary Cox model with a trial-by-treatment interaction and using contrasts to estimate HRs and CIs within each of the trials.

OFS Question: In addition to the overall treatment comparison, the HRs (95% CI) will be estimated within each of the chemotherapy cohorts separately (SOFT prior chemo; SOFT no prior chemo) as described above.

2.6 Analysis Components

2.6.1 CONSORT

For the CONSORT flow diagram. Any changes (except WC/LFU) from first analysis will be noted.

Tables:

- CONSORT diagram content numbers by treatment assignment
    - Number of patients randomized
    - Number of patients included vs excluded from analysis population (noting any difference from first report)
    - Number in analysis population who never started protocol treatment (noting any difference from first report)
    - Number in analysis population who WC/LFU

- Number of patients analyzed in analysis population [same/as number included]

2.6.2 Follow-up compliance and WC/LFU

This section summarizes follow-up compliance, according to trial, in the ITT populations.

Tables:

- Institutional follow-up compliance group/country (rows); by trial and overall
- Withdrawn consent and lost to follow-up status

2.6.2 Stratification factors and patient characteristics

2.6.2.1 Stratification factors

Any patients for whom values of the stratification factors are determined to be different from those originally analyzed will be listed.

2.6.2.2 Characteristics

Because some of the values of characteristics may have changed since the first trial report, tables of characteristics of the analysis populations to be used in subgroup analyses will be re-summarized overall and by trial, cohort and treatment group. Continuous variables are summarized as mean, SD, min/max, and quartiles. Categorical variables are summarized as N(%); for variables with unavailable (missing, unknown, not done) values, the default approach is to list the number of unknowns as a category, but calculate percentages excluding these unknowns from the denominator; however there will be some variables that are frequently not assessed and in those cases the percentages will be calculated with and without the unknown category.

~~Table of:~~

- ~~Patient: age group~~
- ~~Disease: HR status, HER2 status/treatment; N status, T size, T grade~~
- ~~Treatment: local therapy (surgery/RT), chemotherapy regimen~~

2.6.3 Primary efficacy analysis

The primary efficacy analysis will proceed as summarized in Section 2.4 above. The data cut-off and database lock dates used for the analyses and the median follow-up duration will be reported (MFU also by trial & treatment group).

2.6.3.1 Subgroup Analyses and Covariate-adjusted HR Estimates

Same as first analysis.

2.6.3.2 Tables and Figures

Tables:

- AI Question (E+OFS vs T+OFS), stratified by cohort & nodal stratum
  - Primary treatment comparison: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 8yr DFS, SE and 95% CI
  - By-trial (2) and by-cohort (4) treatment comparisons: N events and patients, HR, SE, 95% CI, 8yr DFS, SE and 95% CI

  - Treatment effects within subgroups: N events and patients within each subgroup, treatment HR, 95% CI, 8yr DFS, SE and 95% CI
  - Adjusted model: parameter estimate ± SE, HR, 95% CI for each model covariate
  - TEXT adjusted model (stratified by chemo & nodal strata): parameter estimate ± SE, HR, 95% CI for each model covariate
- OFS Question (T+OFS vs T), stratified by chemo & nodal strata
  - Primary treatment comparison: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 8yr DFS, SE and 95% CI
  - By-chemo cohort (2) treatment comparisons: N events and patients, HR, SE, 95% CI,8yr DFS, SE and 95% CI
  - Treatment effects within subgroups: N events and patients within each subgroup, treatment HR, 95% CI, 8yr DFS, SE and 95% CI
- E+OFS vs T, stratified by chemo & nodal strata
  - Primary treatment comparison: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 8yr DFS, SE and 95% CI
  - By-chemo cohort (2) treatment comparisons: N events and patients, HR, SE, 95% CI,8yr DFS, SE and 95% CI
  - Treatment effects within subgroups: N events and patients within each subgroup, treatment HR, 95% CI, 8yr DFS, SE and 95% CI
- SOFT adjusted model (3-arms; stratified by chemo & nodal strata): parameter estimate ± SE, HR, 95% CI for each model covariate; T as reference group for reporting T+OFS vs T and E+OFS vs T treatment comparisons

Figures:

- AI Question (E+OFS vs T+OFS)
  - KM plot of DFS, by treatment group, for entire analysis population *(y-axis: Percent Alive and Disease-Free; x-axis: Time since Randomization (yearly intervals); x-axis limited to 9 years (min SOFT/TEXT MFU plus 1 yr) and numbers at risk at yearly intervals)*
  - KM plots of DFS, by treatment group, for each trial, and for each of 4 cohorts
  - Forest plot of DFS, overall and for trials & for cohorts
  - Forest plot of DFS, overall and for subgroups

2.6.4 Secondary efficacy analyses

Breast cancer-free interval, distant recurrence-free interval and overall survival will be summarized as described for DFS. Forest plots for BCFI; DRFI and OS depend on # events.

Sites of first treatment failure will be summarized overall and by treatment group, by trial & treatment group, and by cohort(or chemo strata) & treatment group as N (%).

Second (non-breast) malignancies as site of first failure, and deaths without prior cancer event will be summarized overall and by treatment group.

Tables:

- Primary treatment comparison for each of the 3 endpoints: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 8yr DFS, SE and 95% CI

- Cohort treatment comparisons (4) for each of the 3 endpoints: N events and patients, HR, SE, 95% CI,8yr DFS, SE and 95% CI
- Sites of first failure, overall and by treatment group
- Types of second non-breast malignancies, as site of first failure, overall and by treatment group.

Figures:

- KM plots of BCFI, by treatment group, for entire analysis population and for each of the 4 cohorts *(y-axis: Percent Free from Breast Cancer; x-axis)*
- KM plots of DRFI, by treatment group, for entire analysis population and for each of the 4 cohorts *(y-axis: Percent Free from Distant Recurrence)*
- KM plot of OS, by treatment group, for entire analysis population and for each of the 4 cohorts *(y-axis: Percent Alive)*

2.6.5 Adverse Events / Safety

2.6.5.1 AE population

The AE population is the subset of patients in the ITT analysis population who started protocol treatment (and have at least 1 post-baseline AE form submitted). This population will be updated from the first analysis using updated information, as necessary, and any such changes will be included in the report.

2.6.5.2 AE analysis

Targeted AEs, and other grade 3-5 AEs, are collected on CRFs, until 1 year after the cessation of all protocol treatment. The grade and causality attribution are recorded according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) version 3.0. In TEXT, chemotherapy-related AEs are collected separately.

The targeted AEs will be summarized by AE type and maximum grade over time, regardless of causality attribution. The maximum grade consolidates the reports of a given type of AE for a patient over time since randomization (i.e., baseline reports are excluded) by taking the maximum across time (i.e., a patient appears only once for a given type of AE). Patients with reports of multiple AEs of different types are reported multiple times under the relevant AE categories. Maximum grade 0 indicates that the AE type has not been reported.

Note, during a period of trial conduct, the IBCSG audit team was asking sites to record grade 3+ weight gain and obesity as other AEs; discussed among StatC/DMC/CC that CRFs record weight serially on E-form and any analysis of weight changes and/or obesity [in these or other IBCSG adjuvant trials] would use the actual weight, not the AE; as a result the audit team changed their practice. Therefore the weight change and obesity “other grade 3+” AEs will be deleted from dataset before summarizing other grade 3+ AEs because the incidence will not be interpretable.

Tables:

- AI Question (E+OFS vs T+OFS)
    - Targeted AEs reported according to treatment group (N; %)
    - Targeted AEs reported according to trial & treatment group (N;%)
    - Targeted AEs reported according to cohort & treatment group (N;%)

    - Tables above repeated for the subset of events deemed possibly, probably or definitely related to study treatment(s)
- SOFT (E+OFS vs T+OFS vs T)
    - Targeted AEs reported according to treatment group (N; %)
    - Targeted AEs reported according to chemo strata & treatment group (N;%)
    - Tables above repeated for the subset of events deemed possibly, probably or definitely related to study treatment(s)

Other grade 3 or higher AEs are also requested on CRFs, by write-in text. All will be tabulated, similarly according to max grade, but the intention is to focus on those deemed possibly, probably or definitely related to study treatment(s), which will also be tabulated.

2.6.5.3 Post-treatment AEs

The follow-up (E) form collects some events occurring after protocol treatment has ended (which aren't reported on the AE form): clinically-significant cardiovascular and/or cerebrovascular events, bone fractures, and other severe (≥ grade 3) AEs. These will not be part of the update clinical trial report.

2.6.6 Treatment

2.6.6.1 Protocol treatment

Data are from the internal Treatment Summary form; see SDMC Treatment Summary Form Guidelines for details of how data are recorded. There is some inconsistency, relevant primarily for TEXT, among patients who completed 5 years of oral ET with whether any oral ET that continued after 5y since randomization but stopped within year 5-6 was recorded on the protocol treatment page or non-protocol treatment page (i.e., though protocol said 5 yrs since randomization, patients sometimes received 5 yrs since oral ET start; goal is to separate completion of 5 yrs of oral ET from extended adjuvant ET). Guidelines originally said put it all on non-protocol treatment page but this was changed so that up to 6 additional months stayed on protocol treatment page, and if it continued beyond 6 months then protocol treatment ended on day rando+5y and non-protocol treatment started on day rando+5y+1d until actual stop date. So need to be careful in analysis of time until oral ET stop that everything is consistently truncated at 5 yrs.

As of the clinical data cut-off, all patients would have ceased protocol treatment (there are a few patients for whom final confirmation of cessation not received and/or final date of cessation not documented). Protocol treatment status will be summarized, overall and by treatment assignment, and also by trial and/or cohort. The variables are the same as in the original trial report (see original SAP for details), with the exception of the "continuing protocol treatment" category vs. "last known on protocol treatment" categories.

2.6.6.2 Non-protocol treatment

This will not be part of the update clinical trial report. A preliminary analysis will be conducted after compiling the update clinical trial report to get a sense of what is going on; a full analysis of protocol and non-protocol treatment is planned to follow separately). We will plan to describe

- Non-protocol ET to complete 5 years (if protocol-assigned T/E is stopped early);
- [OFS question] Use of OFS within 5 years of randomization among patients assigned tamoxifen-alone

- Concomitant bisphosphonates use
- Extended adjuvant ET use, including oophorectomy/ovarian irradiation after 5 years.

As noted above, among those who completed 5 yrs of protocol-assigned oral ET, need to be careful about an extra few months up to 1 yr of protocol-assigned oral ET, which could be reflected either in the protocol treatment variables or non-protocol treatment variables. And need to be careful not to consider these few extra months as extended adjuvant ET.

Recall, protocol treatment is considered as completed per-protocol at the time of any DFS event. In the case of local recurrence and second (non-breast) cancer, the protocol allowed treatment to continue. The Treatment Summary form non-protocol therapy page captures the therapy after local recurrence and second (non-breast) cancer (as during the trial we clarified that we considered protocol therapy as ended at the DFS event in such cases as well); such therapy is *not* considered as non-protocol therapy for these analyses.

Also in SOFT, if a patient had been on tamoxifen and was randomized to Exe+OFS, the tam was allowed to continue until the exe was started, 6-8 weeks after OFS was initiated; this is also captured on the Treatment Summary form non-protocol therapy page and such therapy is *not* considered as non-protocol therapy for these analyses.

With regard to non-protocol OFS:

- At the time of this database lock, there was no record of non-protocol ovarian irradiation use;
- Any use of GnRH agonist is considered, of any type, and for any duration (as long as not after DFS event);
- Any ooph reported >5yrs for any reason is recorded, as long as it's not after DFS event;
- The timing of non-protocol OFS was segmented as 0 to 4.5yr since randomization vs. ≥4.5yr since randomization; this is in parallel with protocol therapy, for which triptorelin through 4.5yr was considered as completed per protocol.

With regard to non-protocol ET:

- As noted, it can be difficult to differentiate competing 5 yrs of oral ET from extended oral ET use from the 66 month visit; the plan is to identify number of patients on oral ET at ≥6 yrs since rando; look at 2 denominators, ITT, and patients who are DFS event-free & in follow-up at 6 yrs since rando.
- There is just one record of toremifene, which is just being lumped with tamoxifen
-

2.6.6.3 Tables and Figures

Tables:

- Treatment status, overall and by treatment group and trial and/or cohort

Figures:

- Cumulative incidence of early treatment discontinuation, by treatment group, and secondarily by cohort and/or trial and treatment group; without and with competing risk of DFS event that leads to treatment cessation.

2.6.7 GYN and fertility-related issues

During and after protocol treatment, the follow-up (E) form routinely collects menstrual status, pregnancy, and GYN procedures. In addition, after protocol treatment has ended we collect pregnancy attempts. These will not be part of the update clinical trial report. Research projects are planned.