International Breast Cancer Study Group Statistical Center

SOLE (Trial 35-07/BIG 1-07)

Statistical Analysis Plan

Version	Author	Date	Status
1	Weixiu Luo	25 Jan 2017	Approved by …
			Updated reason …

1 INTRODUCTION

1.1 BACKGROUND

In 2006, the standard duration of adjuvant endocrine therapy for breast cancer (either SERMs or AIs) is five years. Patients who receive extended adjuvant letrozole for five years following approximately five years of tamoxifen obtain further benefit compared with the five years of tamoxifen alone. Similarly, benefit has been demonstrated for switching from tamoxifen to an AI after 2 to 3 years of tamoxifen to complete five years of endocrine therapy, as well as initiating therapy with AI following surgery and administering the AI for five years.

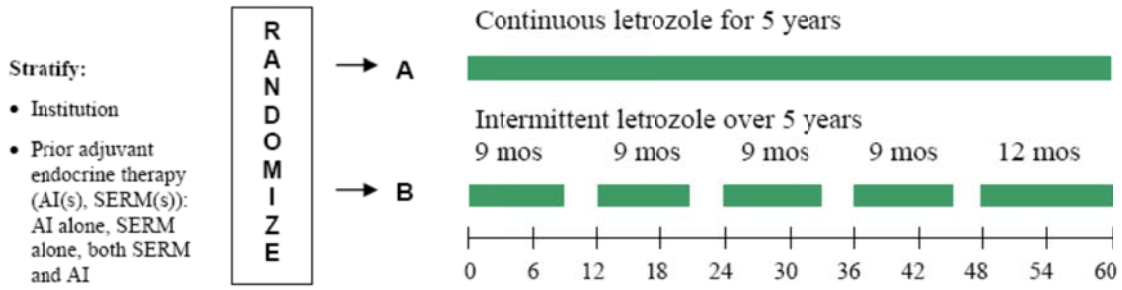
Questions remain about the optimal duration and best schedule of AIs in the extended adjuvant setting. This trial tests the hypothesis that introducing 3-month treatment-free intervals during the course of five years of extended adjuvant letrozole will improve disease-free survival. This hypothesis is based on the theoretical principle that letrozole withdrawal for 3 months will permit some estrogenic stimulation which makes residual resistant disease susceptible to letrozole reintroduction.

1.2 TRIAL DESIGN

SOLE (IBCSG 35-07 / BIG 1-07) OVERVIEW

Title:	**Study Of Letrozole Extension (SOLE):** A phase III trial evaluating the role of continuous letrozole versus intermittent letrozole following 4 to 6 years of prior adjuvant endocrine therapy for postmenopausal women with hormone receptor positive, node-positive early stage breast cancer.
Patient Population:	Postmenopausal women who are disease-free following 4-6 years of prior adjuvant endocrine therapy with selective estrogen receptor modulator(s) (SERM) and/or aromatase inhibitor(s) (AI) for endocrine-responsive, node-positive early breast cancer.
Entry:	Patients must have stopped prior endocrine SERM/AI therapy, and must be randomized within 12 months (1 year) of the last dose of prior endocrine SERM/AI therapy.
Activation Date:	8 November 2007 (First patient randomized on 5 December 2007)
Target Accrual:	4800 patients
Closure Date:	12 July 2012 (Last patient randomized on 8 October 2012)
Final Accrual:	4884 patients

Trial Schema:



The SOLE trial is a multinational phase III randomized clinical trial designed to compare continuous letrozole for 5 years with intermittent letrozole over a 5-year period among postmenopausal women who are disease-free following 4 to 6 years of prior adjuvant endocrine therapy with SERM(s) and/or AI(s) for endocrine-responsive node-positive operable breast cancer. The hypothesis is that introducing 3 month treatment-free intervals during the course of 5 years of extended letrozole will improve disease-free survival.

The randomization was stratified according to participating center and prior SERM/AI endocrine therapy (SERM(s) alone, AI(s) alone, both SERM(s) and AI(s)). Patients were randomized 1:1 to one of two extended adjuvant endocrine therapy groups:

A: Continuous letrozole 2.5 mg daily for 5 years

B: Intermittent letrozole 2.5 mg daily for the first 9 months of years 1 through 4, with treatment-free intervals for the last 3 months of years 1 through 4, followed by 12 months of letrozole 2.5 mg daily in year 5

Patients should start trial treatment as soon as possible after randomization. Trial treatment should begin no later than 6 weeks from the date of randomization. All patients will be followed every 6 months for years 1 to 5, and yearly thereafter for assessment of disease status and for survival data collection.

1.3 SAMPLE SIZE CONSIDERATIONS

The sample size was determined to provide 80% power to detect a 20% reduction in the risk of an event defining DFS associated with intermittent letrozole compared with continuous letrozole (hazard ratio = 0.80; 25% increase in 4-year DFS from 90% to 91.917%) using a two-sided 0.05 level test of significance.

To detect the treatment difference 647 DFS events are required, assuming 4800 patients are accrued (1600 patients per year for 3 years), 5% non-assessability at 4 years, and approximately 5 years of additional follow-up. One year of start-up time, as Participating Centers obtain Ethics Committee approval and complete regulatory processes, is anticipated.

1.4 INTERIM MONITORING PLAN (FROM PROTOCOL)

A group sequential design with two interim analyses and one final analysis is used. The target number of events for the final analysis is 647, and interim analyses are planned after 40% and 70% information (259 and 453 events observed respectively). At each interim analysis and at the final analysis, testing will be performed using O'Brien-Fleming boundaries.

1.5 DATA AND SAFETY MONITORING COMMITTEE REVIEWS

The Data and Safety Monitoring Committee (DSMC) reviews the safety information, including adverse events, second (non-breast) malignancies, and deaths without prior cancer event, every six months and has recommended the trial continue as planned. The first interim efficacy analysis was performed for spring 2014 DSMC report when 324 DFS events (50% information) were observed. At that time the O'Brien-Fleming boundary was not crossed and there was not enough evidence to support early stopping of the trial due to efficacy. The DSMC recommended that SOLE treatment administration and follow-up continue as planned. The same results held for the second interim efficacy analysis which was performed for spring 2015 DSMC report when 454 DFS events (70% information) were observed.

2 EFFICACY ANALYSIS PLANS

2.1 OBJECTIVES

The primary objective is to compare continuous letrozole for five years with intermittent letrozole over a five year period for postmenopausal women who are disease-free following 4-6 years of prior adjuvant endocrine therapy with SERM(s) and/or AI(s) for endocrine-responsive, node-positive, operable breast cancer.

2.2 ANALYSIS POPULATIONS

The primary analysis will use an intention-to-treat (ITT) approach. The ITT population will include all randomized patients, regardless of eligibility status; the possible exceptions are patients who immediately withdrew consent prior to treatment initiation and declined all participation, patients determined (e.g., via audit) to be without documented informed consent, and/or patients at a participating center determined not to be compliant with protocol procedures. Any exclusion from the ITT population will be determined prior to the analysis and will be summarized in listing and CONSORT in the trial report.

2.3 ENDPOINT DEFINITIONS

2.3.1 Primary Endpoint

• Disease-free survival (DFS) is defined as the duration of time from randomization to the first indication of the following events: invasive recurrence at local (including recurrence restricted to the breast after breast conserving treatment), regional or distant sites; a new invasive cancer in the contralateral breast; any secondary (non-breast) malignancy; or a death without prior cancer event. Appearance of DCIS or LCIS either in the ipsilateral or in the contralateral breast will not be considered as an event for DFS. In the absence of an event, DFS is censored at the date of last follow-up.

2.3.2 Secondary Endpoints

• Overall survival (OS) is defined as the duration of time from randomization to death from any cause, or is censored at the date last known alive. (Note, for patients who withdrew consent or were lost to follow-up but follow-up for survival was possible through hospital or registry records, OS is censored at the date last known alive rather than date of last follow-up/withdrawn consent).

• Distant disease-free survival (DDFS) was a secondary endpoint in protocol. It will be replaced by a more modern definition -- distant recurrence-free interval (DRFI) defined as the duration of time from randomization to the first indication of invasive breast recurrence at a distant site. In the absence of an event, DRFI is censored at the date of last follow-up or date of death without distant recurrence.

• Breast cancer-free interval (BCFI) is defined as the duration of time from randomization to the first indication of the following events: invasive breast recurrence at local, regional or distant sites; a new invasive cancer in the contralateral breast. In the absence of an event, BCFI is censored at the date of last follow-up or date of death without prior breast cancer event. (Second non-breast malignancies are ignored)

• Site of First Failure: Hierarchy of failures from least to worst, following standard IBCSG definition:

o Local
o Contralateral Breast
o Regional ± above
o Soft Tissues/Distant Nodes ± above
o Bone ± above
o Viscera (including 'other') ± above

as well as incidence of 2nd (non-breast) malignancy, death without prior cancer event, and adverse events.

For "site of first failure," if there are 2 or more failure sites within 2 months of each other, then the "worst" site is considered as the site of first failure. For example if there is a regional recurrence and a bone recurrence within 2 months of each other, then the bone would be

considered the "worst" site of failure and considered as "site of first failure." If second malignancy is coincident with breast cancer recurrence then site of first failure is considered as the breast cancer recurrence.

Per IBCSG standard, the date of DFS event is the date a proven recurrence was first suspected. In rare cases that the date suspected is prior to randomization, date of randomization is considered as date of event.

2.4 FOLLOW-UP

Number of DFS events per person-year of follow-up is calculated at each DSMC review time to guide the timing for the database lock. The data cut-off is planned for prior to November 1, 2016 when we expect the number of DFS events will reach the target of 647 for the analysis with database lock in Q1 2017. Data as of the database cut-off will be used for 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).

Updated results will be presented every two to three years thereafter.

2.5 TESTS AND ESTIMATES

The primary endpoint, DFS, will be compared between treatment arms using a two-sided stratified logrank test with an overall experiment-wise alpha-level equal to at most 0.05. The test statistic and p-value will be taken from the stratified Cox PH model score test. Hazard ratios will be estimated from a stratified Cox PH model, with 95% CIs. Kaplan-Meier estimates of the DFS distributions will be calculated for each of the treatment arms, with reporting of the 5yr DFS. 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

The stratification factor during randomization is prior SERM/AI endocrine therapy (SERM(s) alone, AI(s) alone, both SERM(s) and AI(s)) as reported on randomization (RA) form. Institutions were balanced using dynamic balancing. Logrank test and Cox PH model will be stratified by prior SERM/AI endocrine therapy.

Note about values of strata as provided at randomization vs. actual values: The randomization form (RA form) collects the values entered into the IBCSG randomization system and used for stratification of the randomization assignment. If these values are incorrect then they are amended on the Registration form (A-form). We will cross-tabulate the stratification variables

between RA and A forms to compare the information obtained at randomization versus that on the A forms. For primary and secondary overall analyses we will use the stratification factor entered on randomization form. For subgroup analyses, we will use the actual values entered in the A form when available.

2.6 ANALYSIS COMPONENTS

2.6.1 Enrollment, Eligibility, Follow-up Compliance

This section summarizes accrual, eligibility, exclusions from ITT population, institutional CRF and follow-up compliance.

2.6.1.1 All randomized patients

Tables:

- Enrollment by group/country (rows), according to year and prior endocrine therapy strata (columns)
- Patient randomization treatment assignment and prior endocrine therapy strata
- Institutional follow-up compliance by group/country (rows)
- Case report form submission status (baseline forms)

Figures:

- Enrollment over time; x-axis time in 6-montly intervals; y-axis number enrolled

2.6.1.2 CONSORT

In the manner of the CONSORT diagram, the following will be summarized.

Tables:

- CONSORT diagram content numbers by treatment assignment
    - Number of patients randomized
    - Number of patients included vs excluded from analysis population, with reasons
        - Listing of patients excluded from analysis population (patid, randomizing institution, randomization date, reason excluded, treatment assignment)
    - Number in analysis population who never started protocol treatment
    - Number in analysis population who WC/LFU
    - Number of patients analyzed in analysis population [same/as number included]

2.6.1.3 Analysis population

Tables:

- Enrollment by group/country (rows)
- Patient randomization assignment and prior endocrine therapy strata

- Eligibility status and reasons ineligible, overall and by treatment assignment
- Listing of ineligible patients (patid, randomizing institution, randomization date, reason ineligible, treatment assignment)
- Withdrawn consent and lost to follow-up status

2.6.2 Patient, Disease and Prior Treatment Characteristics

Characteristics of the analysis population will be summarized overall and by 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 include an unknown category that is included in the denominator for percentages (rather than just listing the number of unknowns as a category).

Tables (overall and by treatment group, unless otherwise specified):

- Patient:
    - Age at randomization (continuous; categorized in 5-year intervals)
    - Race/ethnicity
    - Performance status at randomization
    - BMI at randomization
    - Menstruation status at primary cancer diagnosis
    - Symptoms (from baseline AE form)
- Treatment:
    - Local therapy (combining surgery [Mx/BCS] and radiotherapy [yes/no]
    - Chemotherapy (whether used; regimen [anthracycline-based; taxane-based; both; other]
    - Biologic therapy
    - Prior endocrine therapy and years of prior endocrine therapy
    - Duration from end of prior adjuvant therapy to randomization
- Disease:
    - ER/PgR status and details (H form)
    - HER2 status (H form)
    - Nodal involvement (number of positive lymph nodes 0-1, 2-3, 4-9, and 10+)
    - Tumor size (<1, 1-2, >2-5, >5 cm) and grade (1, 2, 3), primary histology (and other details from H form)
    - Disease laterality and location

2.6.3 Primary Efficacy Analysis

The primary efficacy analysis will proceed as summarized in Section 2.5 above. The data cut-off and database lock dates used for the analyses and the median follow-up duration will be reported.

2.6.3.1 Subgroup Analyses

The protocol pre-specified factors that will be used to characterize the patients enrolled in the study and to provide descriptive statistics of outcomes according to subgroups of the population. These factors include: age at randomization, body mass index, tumor size, tumor grade, number of positive lymph nodes, ER/PgR, HER2 status, type of prior endocrine therapy, duration of prior endocrine therapy by type, interval of time since the cessation of prior endocrine therapy till randomization. These analyses will be considered as secondary and descriptive.

The plans for these variables are summarized below:

- Age (5-year age groups (<55,55-59,60-64,65-69,≥70))
- BMI (obese (>=30), ~~normal or~~ overweight (25 - <30), normal (<25), unknown)
- ~~Chemotherapy (anthracycline-based, taxane-based, both, other, none)~~
- ~~Surgery/RT (Mx+RT; Mx alone; BCS+RT; BCS alone; unknown)~~
- Tumor size (≤2 vs >2 cm; unknown)
- Tumor Grade (1,2,3,unknown)
- ER/PgR subgroup (+/+; +/-; -/+; unknown and other):
- HER2 status (positive, negative, unknown)
- No. of positive lymph nodes (0, 1-3, 4+)
- type of prior endocrine therapy (AI, SERM, AI and SERM)
- duration of prior endocrine therapy (<4.5, 4.5-5.5, >5.5 years)
- Interval of time since the cessation of prior endocrine therapy till randomization (<= 1 month, >1 month)

2.6.3.2 Models

Stratified Cox PH regression models will be used to: estimate HRs (95% CI) for treatment effect, adjusted for these covariates in Section 2.6.3.1; and estimate HRs (95% CI) for treatment effect within subgroups by including treatment-by-covariate interaction in the model (but not other covariates) and using contrasts.

2.6.3.3 Tables and Figures

Tables:

- Primary treatment comparison: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 5yr DFS, SE and 95% CI

- Treatment effects within subgroups: N events and patients within each subgroup, treatment HR, 95% CI, p-value for test of treatment-by-variable interaction

Figures:

- KM plot of DFS, by treatment group, for entire analysis population *(y-axis: Percent Alive and Disease-Free; x-axis: Time since Randomization (12-month intervals); x-axis limited to 6 years (median follow-up plus 1 year) and numbers at risk at each 12-monthly interval)*
- Forest plot of DFS, overall and for subgroups

2.6.4 Secondary Efficacy Endpoints

Breast cancer-free interval, distant recurrence-free interval and overall survival will be summarized as described for DFS. Sites of first treatment failure will be summarized overall and by 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.

2.6.4.1 Tables and Figures

Tables:

- Primary treatment comparison for each of the 3 secondary endpoints: N events and patients, HR, 95% CI, log-rank test statistic and p-value, 5yr 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.
- Death without recurrence (list of patients with cause of death)

Figures:

- KM plots of BCFI, by treatment group, for entire analysis population *(y-axis: Percent Free from Breast Cancer; x-axis: Time since Randomization (12-month intervals); x-axis limited to 6 years (median survival plus 1 year) and numbers at risk at each 12-monthly interval)*
- KM plots of DRFI, by treatment group, for entire analysis population *(y-axis: Percent Free from Distant Recurrence; x-axis: Time since Randomization (12-month intervals); x-axis limited to 6 years (median survival plus 1 year) and numbers at risk at each 12-monthly interval)*
- KM plot of OS, by treatment group, for entire analysis population *(y-axis: Percent Alive; x-axis: Time since Randomization (12-month intervals); x-axis limited to 6 years (median survival plus 1 year) and numbers at risk at each 12-monthly interval)*

2.6.5 Adverse Events / Safety

2.6.5.1 Population

The toxicity population is the subset of patients in the ITT analysis population who started protocol treatment. Any patients without at least 1 post-baseline AE form submitted will not be able to contribute; note any such patients in trial report.

2.6.5.2 Analysis

Targeted AEs, and other grade 3-5 AEs, are collected on CRFs. The grade and causality attribution are recorded according to the NCI Common Terminology Criteria for Adverse Events (CTCAE) version 3.0.

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.

2.6.5.3 Tables and Figures

Tables:

- Targeted AEs reported according to treatment group
- Targeted AEs reported according to prior endocrine therapy and treatment assignment
- 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.

95% exact CIs will be calculated for each targeted AE, according to treatment group.

The following late adverse events (LAEs) will be summarized descriptively in a table according to treatment group: cardiac ischemia/infarction, thrombosis/thrombus/embolism, CNS cerebrovascular ischemia, and fractures.

2.6.6 Treatment

Protocol treatment status as of the clinical data cut-off will be summarized. In this report, status will be summarized by treatment assignment. Time from randomization to treatment termination will be summarized using KM curves according to treatment assignment.

Per SOLE Data Manager's Hand Book v3.0 (dated 21Mar13), for Arm B patients, interruptions should be every 9 months and should last for 3 months in years 1-4. Any interruption up to 4 months is considered a "protocol-specified interruption" and the reason for interruption is defined as "protocol-defined interval" on L forms; any interruption of 4 months +1 day (121 days or more) is considered a "non-protocol interruption".

The Idealized Scheduler contains a link to an Excel file on the IBCSG web site. This file calculates the letrozole dispensation and interruptions (Arm B) based off of the Randomization Date.

Treatment variation scenarios:

- Patient on Arm B stops letrozole early (e.g., at Month 6 instead of Month 9). The patient should stop for three months. Patient should re-start letrozole and continue until the pre-defined date of the next interruption. This will get the patient back on schedule.
- Patient on Arm B took medication during the protocol-specified interruption. Patient should continue until Month 9 of the next year and then start the interruption. This will get the patient back on schedule.
- Patient forgot to take some pills and therefore has some left at the date the interruption should start. The patient should not finish the medication. The patient should interrupt on the pre-defined date to stay on schedule.

For Arm A patients, not-per-protocol interruptions are those of over a month in duration.

Reasons for treatment interruptions will be summarized by treatment assignment. Note that there could be more than one reason for an interruption spanning more than one L Form: for example, a patient start interruption per protocol on one L form, but due to adverse event, the interruption continues to the next L form without restarting after 4 months. The first reason would be "protocol-defined interval" and then the second reason would be "Adverse Event".

Based on information on Data Manager's Handbook, to describe adherence with protocol-assigned treatment, the following tables and figures are proposed.

Table 2.6.6.1. Patients' Treatment Compliance by treatment assignment

	Arm A		Arm B		Overall	
	N	**%**	**N**	**%**	**N**	**%**
Did not start protocol treatment						
Receiving protocol treatment						
Stopped all study treatment						
Completed treatment per protocol						
Stopped all study treatment early						

Adverse event						
Disease related						
Patient decision						
Other						
Patient Lost to Follow-up						
Treatment status unknown due to missing information on forms						
Total number of patients w. treatment data available						

Table 2.6.6.2. Days of Treatment on by treatment assignment during year one to five

	Arm A			Arm B		
Year	N	Median	IQR	N	Median	IQR
1						
2						
3						
4						
5						

Note: N is total number of patients on treatment during that treatment year

Table 2.6.6.3 Number of patients with per-protocol interruptions and resumes for patients on Arm B only

	Per-protocol Interruption			Per-protocol Resume		
Year	n	N	% (n/N)	n	N	% (n/N)
1						
2						
3						
4						
Overall						

Note: N is total number of patients on treatment any time during the treatment year

Table 2.6.6.4 Not-per-protocol interruptions for patients on Arm A and Arm B

	Arm A			Arm B		
Year	n	N	%	n	N	%
1						
2						
3						
4						
5						

Note: N is total number of patients on treatment any time during the treatment year

Table 2.6.6.5 Reasons for treatment interruptions by treatment assignment

	Arm A		Arm B	
Reasons	N	%	N	%
Protocol-defined interval (Arm B only)				
Patient decision				
Medical decision				
adverse event				
Other				

Note: there could be more than one reason for an interruption spanning more than one L Form

Figure 2.6.6.1 Percentage of patients on treatment by month, according to treatment group (x-axis: month 1 to 60, y-axis: percentage)

Figure 2.6.6.2 Percentage of patients on treatment by 3-month intervals, according to treatment group (x-axis: interval 1 to 20, y-axis: percentage)

Figure 2.6.6.3 Mean percentage of days on treatment during 3-month intervals, according to treatment group (x-axis: interval 1 to 20, y-axis: mean percentage)