NCT01761877

Unique Protocol ID: 1RO1 CA1615301A1

Secondary IDs: 12-0080-04, 676847

Study Title: NSAID Effects on Clinical and Imaging Breast Biomarkers

Document Title: Statistical Analysis Plan

Last Modification 11/4/2019

Statistical Analysis Plan

Study Title: Effect of Sulindac on Breast Density in postmenopausal women taking aromatase inhibitors for treatment of their breast cancer.

Study PIs: Alison Stopeck, MD and Patricia Thompson, PhD

Original Study. Our study was originally designed as a randomized, placebo-controlled study of the effect of sulindac (SUL) 150 mg BID on breast density (BD) measured by magnetic resonance imaging (MRD) in postmenopausal women taking adjuvant aromatase inhibitors (AIs). We hypothesized that SUL would lead to a decrease in BD after 12 months and that the decrease would be greater than that observed in the placebo arm. Secondary endpoints included effects of SUL on AI associated arthralgia, quality of life and breast tissue biomarkers. The original accrual targets were 75 in each arm.

Study Modifications. Major cuts to the study budget prior to initiation resulted in a revised study design and modification of the primary objectives. As a result, the primary study goals were modified 1) to obtain evidence for SUL action on BD using MRD measures and 2) to assess the tolerability and safety of SUL in the target population. The placebo arm was dropped, and the study was reduced to a single arm feasibility study based on extensive published evidence for no effect of AI therapy alone for 12 months on BD. At that time, a 4-week washout of non-steroidal anti-inflammatory drugs (NSAIDs) [except low dose aspirin] followed by a pre-intervention observation period of 3 months was introduced to establish the stability of BD measures over time. All planned tissue biomarker studies were dropped. The primary study objectives were modified to assess if SUL at 150 mg BID for 12 months showed evidence for any effect on BD and was well tolerated with secondary emphasis on an effect on blood pressure or arthralgia symptoms. In addition, the study design would allow us to report on: 1) within-patient MRD change over 3 months in the absence of intervention, 2) impact of baseline BD on any change in BD with SUL intervention and 3) the fraction of participants that would undergo repeat needle biopsy procedures.

In consultation with NCI, concern was raised that prior studies of AIs showing no effects of AIs on BD were based on mammography and not the more sensitive MRI method under investigation. As such, an observational (non-randomized) control group comprised of BC patients stable on an AI was included specifically to monitor for AI effect on BD using the MRD method.

Initially, the target sample size was reduced to 75 and 60 in the intervention and observation groups, respectively. A previously planned futility analysis of SUL effect on MRD after 40 participants completed 12 months of intervention was changed to the primary analysis on limited resources and slow accrual. Patients consented to banking tissue biopsies for future biomarker studies.

Modified Study Hypotheses: Because of the modification to the study design, the study hypotheses were modified as follows:

Primary

SUL 150 mg BID for 12 months will decrease MRD in postmenopausal women taking aromatase inhibitor therapy.

a. As a sensitivity analysis of this primary hypothesis, per-study analysis described below will also be performed using only subjects who are adherent with study protocols (adherence is defined as having taken at least 80% of study agent).

b. As a sensitivity analysis of this primary hypothesis, per-study analysis for primary hypothesis will also be performed using only those subjects with baseline BD ≥15%.

Secondary

SUL 150 mg BID for 12 months will be well tolerated and grade 3 adverse events will not exceed 10% and patients who drop out secondary to adverse events will not exceed 10%. *With low sample size of 40, only when there are no grade 3 AE or drop out secondary to AEs, can say it will not exceed 10% (point estimate is 0% with 95% Blyth-Casella-Still CI (0-7.78%).*

Reportable Outcomes

1. Change in BD from baseline at 6 and 12 months by MRD in postmenopausal women receiving AI for treatment of breast cancer and 95% confidence interval will be reported separately for:
    a. participants in the sulindac group who started sulindac and who had at least one MRD measure on intervention,
    b. participants enrolled and started in the observation arm with at least one MRD measure post baseline,
    c. participants adherent to SUL intervention (same caveats as a), and

d. in both arms but reported separately for participants with baseline BD ≥15% (same respective caveats as a and b).

2. For the SUL intervention group, change in BD measured by MRD and 95% confidence interval will be reported for during the 3-month observation period.

3. For the SUL intervention group, change in WOMAC, BPI and FACT-G measures and 95% confidence interval will be reported for participants who *initiated* intervention with SUL 150 mg BID at 3, 6 and 12 months. Separately, change for participants with elevated baseline pain levels for BPI and WOMAC or lower quality of life for FACT-G will be reported. [Specifically, for BPI worst pain, severity, and interference-elevated pain will be defined as a score ≥ 2. For WOMAC, elevated pain will be defined as subscale and total scores above the arm-specific medians will be calculated; for FACT-G a lower quality of life will be defined as subscale and total scores below the median for all participants].

4. For the observation group, change in WOMAC, BPI and FACT-G at 6 and 12 months and 95% confidence interval will be reported. Separately, change for participants with elevated baseline pain levels for BPI and WOMAC or lower quality of life for FACT-G will be reported (as defined above). Secondarily, report on change in individual questions and measurement subscales.
5. For the sulindac group, the proportion of patients agreeing to core needle biopsy at baseline and repeat at 6 months.

Final Data Analysis Plan

Hypothesis Testing: For the final analysis, all baseline characteristics will be summarized with appropriate descriptive statistics.

Linear mixed effects model for longitudinal data will be used for testing the primary hypotheses. The fixed effects of the model will include individual groups (SUL and observation) as well as the combined group (SUL and observation [pooled]) for comparison purposes and each visit (baseline, month 6 and month 12) and for the pooled analysis, an interaction between group and visit. In theory, pooling data for the two groups will allow us to generate a more precise estimate of BD change. *Pooling data together allows us generate a narrower 95% CI of estimated BD change.*

Baseline BD, age, as well as important prognostic variables such BMI will be used as covariates. The model will include a random intercept at the subject level to account for correlations due to the repeated measurements from the same subjects. The dependence structure for longitudinal data from

the same subject will be selected with the Akaike Information Criterion. Log-transformed BD at each time point will be used as the outcome with adjustment for log-transformed baseline BD to improve efficiency and hence power for modeling % change in BD directly.

Note: *There is no prespecified plan to conduct a statistical test comparing change in BD between the two study groups.*

Only subjects who start SUL intervention will be included in the primary analysis. This allows us to evaluate the effect of SUL in those individuals who start intervention. 95% confidence intervals of BD change at month 6 and month 12 in the SUL group will be reported.

There are few patients whose MRD could not be obtained because of image artifact. As a sensitivity analysis, multiple imputation will be used to impute all missing MRD using patients' baseline information such as age, race, BMI, cancer stage, chemotherapy and radiation therapy information, AI use and NSAID use. Ten imputed datasets will be generated and all estimates about BD% change will be constructed using Rubin's rule. All analysis for primary hypothesis will be repeated among a subgroup of patients whose baseline BD ≥ 15% and further using only SUL patients with >80% drug adherence.

For secondary hypothesis, all adverse events and severe adverse events from patients in the SUL group will be reported. The rate of patients with grade 3 or higher adverse events will be reported along with 95% Blyth-Casella-Still confidence interval.

For Reportable Outcomes

The linear mixed effect model fitted for the primary analysis will also be used to estimate the BD% changes and 95% confidence intervals at month 6 and month 12 within the observational group.

As part of this analysis for reporting purposes, we will describe the comparability of the intervention and non-random observation arm, continuous characteristics will be compared among study arms using Welch's two-sample t-test. Categorical variables will be compared using Chi-square test and exact p-values based on simulations may be reported if there is data sparsity. Any differences will be reported.

Percentage change in BD during the 3-month observation period in the SUL group will be reported with its 95% confidence interval (*i.e.* % change as [BD at the end of observation period – BD at the beginning of observation period]/BD at the beginning of observation period).

Linear mixed models similar with the one used for primary hypothesis will be carried out to estimate the change in pain (WOMAC and BPI) and quality of life (FACT-G) and 95% confidence intervals separately for the two groups of the study. Three subscales in WOMAC will also be assessed. FACT-G has 4 subscales plus a total score, and BPI has 2 subscales plus the single-question "worst pain" measure. Log-transformation or other type of data transformation may be used to make the model assumption met. No formal comparison between groups will be conducted.

The rate of participants enrolled in SUL intervention that will undergo research breast biopsy and the rate of those who agree to undergo a second research biopsy at 6 month will be reported along with their corresponding 95% Blyth-Casella-Still confidence interval.

New Exploratory Aim.

During the conduct of the trial, recent evidence has emerged regarding the role of collagen as a determinant of BD with evidence for a role for collagen alignment in breast cancer risk. This led to the exploratory hypothesis that any effect of SUL on breast density may be related to effects on collagen alignment measured by second harmonics imaging of tissue biopsy samples. We will also explore the effect of SUL on breast tissue collagen alignment to include effects on collagen straightness, fiber length, width and curliness using similar linear mixed models to estimate change in these measures. The linear correlation between the change in MRD and the change in each measure from tissue biopsy imaging will be assessed using Pearson's correlation coefficient or Spearman's rank correlation coefficient if more appropriate.