Statistical Analysis Plan
Final Analysis

A PHASE III RANDOMIZED TRIAL OF METFORMIN VERSUS PLACEBO ON RECURRENCE AND SURVIVAL IN EARLY STAGE BREAST CANCER

CCTG Protocol Number: MA.32

Prepared by:

Protocol Statistician: Bingshu Chen

Reviewed by:

Senior Investigator: Wendy Parulekar
Study Chair: Pam Goodwin

Date: December 9, 2021

Table of Contents

1 Introduction:

MA.32 is a phase III superiority trial comparing invasive disease free survival (IDFS) between subjects treated with metformin (850 mg po bid for 5 years) versus placebo in addition to standard adjuvant therapy in women with early stage breast caner. This document is to describe the statistical analysis plan for the protocol specified final analysis of MA.32.

1.1 Objectives

The Primary Objective of the trial is to compare improvement in IDFS between treatment and placebo arms in the hormone receptor (ER and PgR) positive breast cancer patients. The Secondary Objectives include comparing the following endpoints: Overall Survival (OS), distant disease-free survival, breast cancer free interval (BCFI), Breast Cancer Specific Mortality (BCSM), contralateral breast cancer (Invasive and DCIS), changes in body mass index (BMI = weight (kg)/height(m)^2), adverse events, other medical endpoints – including a new diagnosis of diabetes mellitus or cardiovascular hospitalization or death, health related quality of life, plasma metabolic factors (including insulin, glucose, CRP, leptin and others) and molecular markers (including insulin receptor, pAKT and others) of metformin action, insulin resistance syndrome.

1.2 Sample Size Determination

With an overall two-sided alpha of 0.05 and 80% power, a total of 554 IDFS events would allow detection of a HR 0.785, after taking into account two interim analyses (see Section 14.5 of the protocol for details). A HR of 0.785 represents an absolute improvement of 3.7% in 5 year IDFS from 81% to 84.7% for those allocated to metformin. The original sample size estimation was 3582 women, assuming 3 years of accrual and an additional 3 years of follow-up, 3582 women will be accrued. A total of 3649 women were accrued to enable international participation The study population for the primary analysis was amended to include to the ER/PgR positive patients after the second interim analysis. See 1.3 for details.

1.3 Timing of the Analyses

Original timing of the analysis

Two interim analyses were planned for this study when 185 and 370 events are observed, to allow early termination of the study if the results were extreme. Lan-DeMets error spending function was used to assess for superiority, and futility for superiority. The early stopping boundaries werebased on a power family with power 3, which approximates the O'Brien-Flemming boundaries. The actual p-values for superiority and futility were calculated based on the number of events observed at the time of interim analysis, controlling the two-sided Type I error of 0.05 and the power of 80% at the end of the study.

If exactly 185 events were observed for the first interim analysis, the null hypothesis would be rejected early due to evidence of superiority at the first interim analysis if the p-value is less than or equal to 0.00185. The alternative hypothesis would be rejected for futility if the p-value was at least 0.971. If exactly 370 events were observed for the second interim analysis, the nominal critical p-values for rejecting the null hypothesis and the alternative hypothesis would have been 0.0138 and 0.468,

respectively. The nominal significant level for the final analysis would be 0.0463 when 554 events were observed, to maintain the overall two-sided alpha of 0.05.

Revised timing of the analysis

Based on the recommendation of DSMC after the second interim analysis, the final analysis would include only study subjects who were ER and/or PgR+ (measured at baseline) and be performed when 554 events were observed from these subjects. Because of the alpha spending during the first two interim analyses, the nominal two-sided significance level for the final analysis would be **0.037** for the revised study population. With 554 events in the final analysis, there would be 78% power to detect the hazard ratio of 0.785 at two-sided 3.7% level or 80% power to detect a hazard ratio of 0.78. It was expected that additional 4 years of follow-up from November 2015 would be required to observe the required number of events. The power would be reduced if the number of events did not reach 554.

AS of November 5, 2020, there were 446 IDFS events in the ER and/or PgR piositive group and based on the current event rate, it is projected that at least another 3 years of follow-up would be required to observe 554 events. With 446 events, the power to detect the hazard ratio of 0.78 would be 70%. The power to detect an hazard ratio of 0.757 would be 80%. Because of the relatively moderate loss in the power (from 80% to 70%) of the study, continued decline in the event rate (from 40 per year in 2018, 27 per year in 2019 and 13 events in the first 10 months of 2020), the Trial Leadership recommended that a time based final analysis be performed in 2021. This request was was approved by the Cancer Therapy Evaluation Program, Central Institutional Review Board with notification and approval by the CCTG independent Data Safety Monitoring Committee. The estimated median follow-up of all women at the time of final analysis would be be approximately 7.6 years.

The selected clinical cut-off date is October 31st, 2020 and the database was cleaned and locked on July 9, 2021.

1.4 Data Collection

Data were collected, entered and managed by CCTG, Queen's University, according to the group standard data management procedures.

2 Methods and Analyses

2.1 Analyses Samples

Only study subjects who were ER and/or PgR+ (variable measured at baseline) will be included in this primary final analysis. There will be parallel analyses in hormone receptor negative subjects and in all subjects combined as secondary analyses. The study populations for this analysis will include both the intention to treat (ITT) and as treated populations with data included as specified by the data cutoff date.

Analysis of pretreatment characteristics and all efficacy outcomes such as IDFS and OS, will be based on the ITT population, regardless of actual treatment received. An sensitivity analysis for efficacy outcomes

will be conducted for the as treated population. Those who received at least one dose of protocol therapy (i.e. the as treated population) will form the basis of the safety analyses.

We will conduct secondary analyses in ER and PgR negative population and the full study population. The results for these two analyses will be provided in a separatestatistical snalysis report (SAR).

2.2 Conventions for Calculating Key Data

In general, baseline evaluations are those collected closest, but prior to or on the day of randomization. If pre-randomization assessment was not done, a pre-treatment assessment will be used as baseline assessment.

When either day or month of a date is missing, the missing day and/or month will be imputed by the midpoints within the smallest known interval. For example, if the day of the month is missing for any date used in a calculation, the 15th of the month will be used to replace the missing day. If the month and day of the year are missing for any date used in a calculation, the first of July of the corresponding year will be used to replace the missing data.

2.3 Analysis Conventions

All comparisons between treatment arms will be carried out using a two-sided test at an alpha level of 5% unless otherwise specified. For the final analysis of the primary endpoint IDFS, the p-value cutpoint will be 0.037 instead of 0.05. No formal adjustments will be made for the multiplicity of inferences for the other clinical endpoints.

The following baseline stratification factors that will be used to adjust the analyses where appropriate are listed below:

- BMI </= 30 versus > 30 (kg/m2)
- HER2 positive versus HER2 negative
- Chemotherapy – any versus none

*Add missing/unknown category whenever appropriate.

2.4 Randomization and Pre-treatment Characteristics

2.4.1 Definitions and Variables

2.4.1.1 Accrual

- Number (%) of randomized patients per study center and country (table 1).

2.4.1.2 Randomization/Stratification

- BMI </= 30 versus > 30 (kg/m2)
- HER2 positive versus HER2 negative
- Chemotherapy – any versus none

A minimization procedure for treatment assignment was used in this study. Stratification factors at randomization will be summarized by treatment arm (table 2).

Number and percentage of actual treatment received will be summaried by treatment arm (table 3).

Treatment at randomization will be compared with the actual treatment received to identify any discrepancies (table 4).

2.4.1.3 Ineligibility and Major Protocol Violations

Number and percentage of ineligible patients will be be presented by treatment arm (table 5).

Reasons for ineligibility: percentage for each reason and combination of reasons of ineligibility will be presented by treatment arm.

The number and percentage of major protocol violations will be presented by treatment arm. (table 5)

The number and percentage of COVID protocol variances will be presented by treatment arm. (table 5a)

2.4.1.4 Summary of Follow-up

A table showing the median, min and max follow-up (defined as reverse censoring on IDFS survival) will be presented by treatment group and for all patients included in analysis. (table 6)

2.4.1.5 Patient Characteristics

Patient characteristics at baseline are summarized in table 7.

- Age
- Height (cm)
- Weight (kg)
- BMI
- Race
- ECOG PS
- Menopausal Status
- Receipt of (neo)adjuvant chemotherapy
- Receipt of (neo)adjuvant hormone therapy
- Hormone receptor status (must be 100% positive)
- cT Stage (neoadjuvant)
- cN Stage (neoadjvuant)

- pT stage (adjuvant)
- pN stage (adjuvant)
- Most extensive primary surgery – breast and axilla separately
- Adj radiotherapy
- HER2 status
- Grade
- Breast cancer sub-type (Ductal vs Lobular)
- Resection margins
- Number of axillary nodes examined
- Number of positive axillary nodes
- Number of sentinel nodes examined
- Number of positive sentinel nodes

*An unknown category will be added when appropriate.

2.4.1.6 Baseline Cancer Treatment

Number and percent of patients who received cancer treatment (Chemotherapy, Hormone Therapy (Neoadj vs post op), Immunotherapy, Adjuvant therapy, Neo-adjuvant therapy, Radiotherapy, Other therapy), Comorbidity (cardiovascular) will be summarized by treatment arm (table 7).

2.4.1.7 Baseline Hematology/Biochemistry

CTC 4.0 grades will be used to summarize the baseline hematology/biochemistry data (% with each CTC grades) for the following assessments (table 8):

- WBC,
- Granulocytes,
- Platelets,
- Hemoglobin

2.4.1.8 Baseline Non-Hematologic Adverse Events

CTC 4.0 grades will be used to summarize the baseline Non-Hematologic (number for each CTC grades, total number and %) (table 9).

2.4.2 Analysis of pre-treatment characteristics

No formal statistical tests will be performed to assess the homogeneity of baseline characteristics between the arms. Categorical variables will be tabulated by treatment arm and for all patients. Continuous variables (e.g., age) or transformed continuous variables (e.g. insulin level) will be presented

using summary statistics (n, mean, standard deviation (SD), median, min and max) or specified cutoff categories by treatment arm and for all patients. Analyses will be based on all randomized patients by arm based on the ITT population.

2.5 Efficacy

2.5.1 Definitions and Variables

2.5.1.1 Invasive Disease-free Survival (IDSF)

The primary endpoint of this study is invasive disease-free survival. It is defined as the time in months from randomization to the time of documented ipsilateral, contralateral invasive breast tumour, local/regional invasive recurrence, distant recurrence, death from breast cancer, death from non-breast cancer cause, death from unknown cause, second primary invasive cancer (non-breast, except for adequately treated BCC or SCC of the skin). If a subject has not had invasive disease nor died at the time of data cut-off for this analysis, IDFS will be censored on the date of last disease assessment.

IDFS Time (months) =
((date of IDFS event or last disease assessment date - date of randomization) + 1) / 30.4375

2.5.1.2 Overall Survival (OS)

For patients who have died, overall survival is calculated in months from the day of randomization to date of death. Otherwise, survival is censored at the last day the patient is known alive (LKA).

OS Time (months) =
((date of death or LKA - date of randomization) + 1) / 30.4375

2.5.1.3 Breast Cancer Free Interval (BCFI)

Breast Cancer Free Interval is defined as the time from randomization to the time of diagnosis of invasive ipsilateral breast tumour recurrence, local/regional invasive recurrence, distant recurrence, death from breast cancer, invasive contralateral breast cancer, ipsilateral DCIS, contralateral DCIS. If a subject has not had BCFI event at the time of data cut-off for this analysis, or died from other reasons, the BCFI will be censored on the date of last disease assessment.

Incidence of contralateral invasive breast cancer is defined as the diagnosis of a primary invasive breast cancer in the opposite breast after randomization.

Number of the contralateral breast cancer (invasive and invasive / DCIS) will be reported by treatment arms.

Types of BCFI events will be listed by treatment arms.

2.5.1.4 Distant Relapse Free Survival (DRFS)

Distant Relapse Free Survival (DRFS) is defined as the time from randomization to the time of distant recurrence, death from breast cancer, death from a non breast cancer cause or death from an unknown cause. If a subject has not had distant RFS event nor died at the time of data cut-off for this analysis, DRFS will be censored on the date of last disease assessment.

2.5.1.6 Breast Cancer Specific Mortalityl (BCSM)

Breast cancer specific mortality is defined as the time from randomization to the time of death from breast cancer. Patients died from causes other than breast cancer will be treated as a competing event and patients who were alive will be censored at the last day the patient is known alive (LKA).

2.5.1.7 IBCFS

IBCFS is defined as the time in months from randomization to the time of documented local/regional invasive recurrence, distant recurrence, ipsilateral and contralateral invasive breast tumour, death from breast cancer, death from other non-breast cancer cause, death from unknown cause. If a subject has not had IBCFS event nor died at the time of data cut-off for this analysis, IBCFS will be censored on the date of last disease assessment.

2.5.1.7 Medical events

1. Diabetes: initiation of new anti-diabetes medication (or confirmed MD diagnosis of diabetes)
2. Cardiovascular hospitalization or cardiovascular death (stroke, myocardial infarction)

2.5.1.9 Summary of events

First Event	IDFS	BCFI	DRFS	BCSM	OS	IBCFS
Local or Regional Recurrence	Event	Event				Event
Distant Recurrence	Event	Event	Event			Event
New primary malignancy	Event					
Invasive ipsilateral breast tumour	Event	Event				Event
DCIS ipsilateral breast tumour(data in F5, F5s)		Event				
Invasive contralateral breast tumour	Event	Event				Event
DCIS contralateral breast tumour (data in F5, F5s)		Event				
Death (Breast Cancer)	Event	Event	Event	Event	Event	Event
Death (Protocol treatment of the disease)	Event	Event	Event	Event	Event	Event
Death (Non-Protocol treatment of the disease)	Event	Event	Event	Event	Event	Event
Death (Others)	Event	Censor	Event	Competing Event	Event	Event
Death (Unknown)	Event	Censor	Event	Competing Event	Event	Event
None	Censor	Censor	Censor	Censor	Censor	Censor

2.5.2 Analysis of Key Parameters

The comparison between treatment arms will be carried out using a two-sided test at an alpha level of 5% unless otherwise specified (e.g. 3.7% for IDFS). All efficacy analyses will be presented by treatment arm. The CONSORT diagram will be included.

2.5.2.1 Invasive Disease-free Survival

A Kaplan-Meier curve for IDFS in each treatment arm will be displayed. The difference between the two treatment arms will be tested using the log-rank test stratified by BMI </= 30 versus > 30 (kg/m2), HER2 positive versus HER2 negative, Chemotherapy – any versus none.

The analysis of IDFS will also be presented for each level of each stratification factor. In addition, as an exploratory analysis, a stratified Cox regression model adjusting for other prognostic factors which are unbalanced or correlated to the IDFS in univariate analyses will be applied to verify the impact of the

prognostic factors on the treatment effect and/or identify factors significantly related to the IDFS. In these analyses, the final Cox model will be determined using a stepwise model building procedure with a significance level of 0.1 used for entry or removal of variables. A un-stratified Cox regression model with stratification factors as covariates will also be fitted. In all regression analyses, all patients with missing value in any covariate will be either excluded from the analysis or imputed to the mean value of that covariate. (Table 10 11, 12, Figure 1).

For patients who developed IDFS events, the type of progression events and summary of progression will be tabulated (Table 12). The percentage of patients with new primary malignancy for each arm will be reported.

A separate Statistical Analysis Plan is developed for blood variables.

2.5.2.2 Overall Survival, BCFI and distant RFS, BCSM IBCFS

The analyses of OS, BCFI, distant RFS, BCSM, IBCFS will be similar to these of IDFS (table 13, 14, 15, Figure 2, 3, 4 will be repeated).

Cumulative incidence curves will be plotted by treastment arms for breast cancer specific survival (BCSS) with death from any causes treated as a competing event.

Medical outcomes (cardiovascular, diabetes) will be summary by treatment arm and compared using a Chi-squared test.

Sensitivity analyses:

- Sensitivity analyses for both IDFS and OS BCFI, distand RFS, BCSS other medical conditions will be performed with the as treated population.

2.6 Drug Exposure

Drug duration

Duration of drug exposure will be reported as time from date of first dose of protocol therapy to the time of last drug dose before the clinical cut-off date. Date of randomization will be used if first dose date is unknown. Patients with IDFS event will be censored at the date of IDFS event for the purpose of drug exposure duration. If date of discontinuation is unknown we use drug dispensed.

Min, Median (based on KM method, censored at IDFS date), and Max duration of drug exposure will be reported by treatment arm in Table 16a.

Drug duration by SNP status

Drug duration: (off protocol treatment) in relation to SNP status will be reported by treatment arms in the subset of patients with available data (Table 16b).

- Comparison1: CC or AC vs CC
- Comparison2: CC vs CA vs AA

Drug adherence

Overall drug adherence (compliance)

The overall compliance rate will be calculated for each patient individually then take an average of all patients in each arm.

For each patient, the overall drug compliance rate =

(Total number of kits dispense * 400 - total number of pills returned) divided by (total number of days on protocol treatment with full doses * 2 pills per day + 1 pill per day with dose reduction)

Total number of days on protocol therapy = days from randomization to first date of the following event:

(IDFS event, off treatment, or withdrawal of consent).

Note: If a kit was dispensed and was not returned, then the number of returned pills count will be estimated at 400. If a patient was off protocol treatment before IDFS event, then the patient will be designated as non-compliant and the drug compliance rate for this patient will be 0%. Numbers and percent of non-compliants will be reported for each arm (Table 16c1).

Drug compliance rates over each time priod

In each study period (month 6, years 1, 2, 3, 4, 5) before iDFS event, off protocol treatment or widthdraw of consent, the drug adherence rate without adjust for dose reduction is calculated as

Drug adherence rate for each drug kit (%) =
(Total number of pills dispensed – Total number of pills returned) / Total number pills dispensed
= (400 – Total number of pills returned)/(400-34.75)

Note: 1) The total number of pills dispense is 400 pills every 6 months and each patient need 365.25 pills on average for 6 months on average (2 pills per day).
2) Some patients returned empty drug kit with will result a drug compliance for a particular kit to be greater than 100% (400/365.25 = 109.51%). According to the request from Dawn, this number will be capped at 100%.

All kits dispensed between momths 0 to 3 will be considered as drug for months 0 to 6.
All kits dispensed between momths 4 to 9 will be considered as as drug for months 7 to 12.
All kits dispensed between momths 10 to 21 will be considered as drug for as months 13 to 24.
All kits dispensed between momths 22 to 33 will be considered as as drug for months 25 to 36.

All kits dispensed between momths 34 to 45 will be considered as as drug for months 37 to 48.
All kits dispensed between momths 46 to 66 will be considered as as drug for months 49 to 60.

The overall adherence without adjust for dose modification rate and the average drug kit compliance rate each time period compliance rates and will be reported by treatment arm (Table 16c2). A sensivity analysis will be conducted with the assumption that the missing kit was 50% and 100% taken.

Dose Modification

Numbers and percent of patients with at least one drug modification (Yes vs No) will be report by each study period (month 6, years 1, 2, 3, 4, 5) and trearment arm (Table 16d).

2.7 Safety

2.7.1 Definitions and variables

All toxicity/side effects data collected post randomization will be included in the analyses of toxicities.

2.7.1.1 Non-hematologic adverse events

Non-hematologic adverse events will be summarized according to CTC AE 4.0 (table 19).

Sub group analysis of non-hematologic AE will be reprted by SNP status(Four more tables for subsets: SNP_base = 'CC', SNP_base = 'AC', SNP_base = 'AA', and SNP_base in ('CC', 'CA')).

- Comparison 1: CC or AC vs AA
- Comparison 2: CC vs AC vs AA

2.7.1.2 Hematology/Biochemistry

Hematology/Biochemistry experienced will be reported according to CTC AE 4.0 (table 20).

2.7.1.3 Serious adverse event

SAE will be listed by treatment arm (table 21).

2.8 Off protocol therapy and death

Patients off protocol treatment): Number and % of all treated patients. Reason for going off protocol therapy (e.g. adverse event, death, progression, etc): Number and % of all treated patients will be presented (table 22).

Deaths within 30 days from last treatment administration. Cause of death within 30 days from last treatment administration: Number and % of all treated patients will be presented (table 23).

2.9 Quality of Life

The EORTC QLQ-30 Global Score will be used for our primary assessment of quality of life but subscales and specific symptoms (diarrhea, bloating, flatulence, dyspepsia, abdominal cramps, nausea and vomiting, taste alteration, limitation of activities because of gastrointestinal symptoms, joint/musculoskeletal symptoms) will be used for our secondary hypothesis.

2.9.1 Definitions and Variables

2.9.1.1. EORTC QLQ-C30

There are five functional domains and three symptom domains that can be derived from EORTC QLQ-C30 (see below for definitions). If the number of unanswered questions in each domain is within a limit specified with the definition for each domain, the score is calculated as below:

For function domains of Physical, Emotional, Cognitive and Social:

Score=100-(((Total score for the answered questions/(Total questions answered))-1)*100/3).

For symptom domains:

Score=(((Total for the answered questions/(Total questions answered))-1)*100/3)

Otherwise, the score will be recorded as "missing". For each single item, the score will be recorded as "missing" if the answer to this item is missing.

Functional Domains:

Physical: Questions: 1, 2, 3, 4, 5, 6, 7
Score=missing if number of above questions not answered is greater than 3;
Emotional: Questions: 21, 22, 23, 24
Score=missing if number of above questions not answered is greater than 2;
Cognitive: Questions: 20, 25
Score=missing if number of above questions not answered is greater than 0;
Social: Questions: 26, 27
Score=missing if number of above questions not answered is greater than 0;

Symptom Domains:

Fatigue: Questions: 10, 12, 18
Score=missing if number of above questions not answered is greater than 1;
Nausea and vomiting: Questions: 14, 15
Score=missing if number of above questions not answered is greater than 0;
Pain: Questions: 9, 19
Score=missing if number of above questions not answered is greater than 0.

There are also six single items in EORTC QLQ-C30 pertaining to common symptoms and one global assessment that can be derived from EORTC QLQ-C30. The single items are:

Single Items:

Dyspnea:	Question 8;
Sleep/Insomnia:	Question 11;
Appetite:	Question 13;
Constipation:	Question 16;
Diarrhea:	Question 17;
Financial:	Question 28.

They are all scored using the following formula:

Score= (Answer to the question-1)*100/3.

The Global Assessment includes Questions 29 and 30. If number of the questions not answered is greater than 0, its score will be "missing"; Otherwise,

Score= ((Total for the answered questions/(Total questions answered))-1)*100/6.

For functional domains, a higher score indicates better functioning but for symptoms domains, a higher scores indicate higher (worse) symptoms.

Trial specific checklist:

The following trial specific checklist were collected:

31. Did you have bloating of your abdomen (stomach or belly)?
32. Did you have pain or cramps in your abdomen (stomach or belly)?
33. Did you have heartburn?
34. Did you have gas?
35. Did you limit your activities because of gastro-intestinal problems?
36. Did you have a metallic taste?
37. Do you have aching muscles or joints?

Exploaraory Analysis

QLQ-C30 Summary Score =

(PF + RF + SF + EF + CF + 100-Fatigue + 100-Pain + 100-Nausea/Vomiting + 100-Dyspnoea + 100-Sleeping Disturbances + 100 Appetite Loss + 100-Constipation + 100-Diarrhoea)/13

Only calculated if all 13 scores available; Scale scores based on completed items (at least 50% completed)

A frequency table to to answer of each question will be listed at baseline and each of the follow-up time.

2.9.2 Analysis

All analyses on quality of life scores will be exploratory and will **include all randomized patients with QoL data**. Report number included in the QOL substudy (888).

The patient characteristics between QOL participants with those who did not participate will be reported. Table 7 will be repeated for patients in QoL study vs patients did not in the QoL study.

Repeat Table 7 again for all Canadian population.

The patient characteristics (Table 7) among the following three groups: no endocrine therapy, tamoxifen and aromatase inhibitor will also be reported.

The time point of the primary QOL endpoint will be 12 months after randomization.

2.9.2.1 Determination of Assessment Times

The following will be the scheme to determining the time frame of a QOL assessment:

1) Baseline: Baseline evaluation is the QOL questionnaire collected closest, but prior to, the first day of starting study treatment/randomization;
2) Monnth 6 evaluation: If the QOL is assessed between month 3 and month 9
3) Monnth 12 evaluation: If the QOL is assessed between month 10 and month 18
4) Monnth 24 evaluation: If the QOL is assessed between month 19 and month 30
5) Monnth 36 evaluation: If the QOL is assessed between month 31 and month 42
6) Monnth 48 evaluation: If the QOL is assessed between month 43 and month 54
7) Monnth 60 evaluation: If the QOL is assessed between month 55 and month 66

2.9.2.2 Calculation of Compliance Rates

The following method will be used to calculate the compliance rates of QOL assessment. The compliance rate is calculated as the number of forms received out of the number of forms expected at each assessment point defined based on the following principles:

1) At baseline: the number of forms expected is the total number of patients who are eligible for the study and required to fill out QOL questionnaires.
2) FU period: the number expected at each assessment is the number of patients with baseline data minus the number of patients who have off the protocol treatment, died or progressed during that and previous follow up period (with assessment window defined by 2.9.2.1).

2.9.2.3 Cross-sectional analysis

The mean and standard deviation of QOL scores at baseline and mean and standard deviation of QOL change scores from baseline at each assessment time will be calculated. Then Wilcoxon Rank-Sum test is

used to compare two treatment arms in terms of change in QOL score at each assessment time from baseline.

Mean change in scores over time will be analyzed using generalized linear mixed model. The treatment arms, patient substudy groups defined by adjuvant endocrine therapy use at the time of randomization (tamoxifen, aromatase inhibitor, no endocrine therapy), and qol assessment time point will be used as covariates in this analysis. The interactions between treatment arms and qol assessment time (up to 60 months) will be tested using a likelihood ratio method.

2.9.2.4 QOL response analysis

QOL response is calculated as follows for a functional domain: A change score of 10 points from baseline was defined as clinically relevant. Patients will be assessed as improved if they have reported a score of 10-points or better than baseline at any time of the QOL assessment. Conversely, patients will be assessed as worsened if there is a reported score that is at least 10 points worse than baseline at any time of the QOL assessments without meeting the criteria for improved. Patients whose scores are intermediate between these values at every QOL assessment will be considered as stable. In contrast to functional domains, for the determination of patient's QOL response, classification of patients into improved and worsened categories is reversed for symptom domains and single items. A Chi-square test will be performed to compare the distributions of these three categories between two arms (improved, stable or worse).

2.9.2.5 QOL subscale trend analysis

Plots of QoL score over time by treatment arms over time (From baseline to month 60) will be reported for gloval QOL, each subscales and symptoms items.

Note: Physical Activity Items of the Nurses Health Study II Questionnaire and Block Alive Screener (Diet) will be analyzed in a saperate analysis report after the finish of the final analysis of MA.32. The levels of fatigue and overall QoL will be evaluated according to diet and physical activity in the metformin and placebo groups in the subset of QoL participants with diet and physical activity assessment.

3 Tables

Table 1 Accrual by centre

Centre	Number of accrual (%)		
	Metformin N = ***	Placebo N = ***	Total N = ***
xxxx	XX (XX)	XX (XX)	XX (XX)

Table 2: Accrual by Stratification Factors at Randomization

Data set: All Randomized Patients			
	Number of patients (%)		
	Metformin N = ***	Placebo N = ***	Total N = ***
ER and/or PgR			
Positive	** (**)	** (**)	** (**)
Both Negative	** (**)	** (**)	** (**)
BMI			
</=30			
>30	** (**)	** (**)	** (**)
	** (**)	** (**)	** (**)
HER2			
Positive			
Negative			
Chemotherapy			
Any			
None			
Source: Centralized Randomization File			

Table 3: Treatment received

	Number of accrual (%)		
Actual treatment received	Metformin N = ***	Placebo N = ***	Total N = ***
Metformin	XX (XX)	XX (XX)	XX (XX)
Placebo			
No treatment			

Table 4: Stratifiation factor at randomization vs. at baseline

Data set: All Randomized Patients			
	Number of patients (%)		
At randomization	At baseline	Metformin N = ***	Placebo N = ***
BMI		** (**)	** (**)
</=30	</=30	** (**)	** (**)
</=30	>30		
>30	</=30		
>30	>30		
HER2			
Positive	Positive		
Positive	Negative		
Negative	Positive		
Negative	Negative		
Chemotherapy			
Any	Any		
Any	None		
None	Any		
None	None		

Table 5: Eligibility status

Total patients allocated	Metformin N = ***	Placebo N = ***	Total N = ***
Ineligible	XX (XX)		
Total eligible patients	XX (XX)		
REASONS FOR INELIGIBILITY			
Reason 1	XX (XX)		
Reason 2	XX (XX)		
Major Protocol Violations			
XXX			
XXX			

Table 5a: COVID-19 Protocol variance

Total patients allocated	Metformin N = ***	Placebo N = ***	Total N = ***
Protocol variance due to COVID19			
Alternate patient assessment (Phone or virtual visit)			
Missed study Visit			
Delayed response assessment			
Other			

Table 6: Summary of Follow-up

Data set: All Randomized Patients			
	Number of patients		
	Placebo N = ***	Metformin N = ***	Total N = ***
Median*	***	***	***
Min	**	**	**
Max	**	**	**

Median Follow-up based on inverse OS

Table 7: Patient characteristics

All randomized patients	Metformin N = ***	Placebo N = ***	Total N = ***
AGE			
<=39	XX (XX)		
40-49	XX (XX)		
50-59			
60-69			
>=70	XX (XX)		
Median (Range)	XX (XX)		
Gender			
M			
F			
Race			
Asian			
Black or African American			
etc			
ECOG PS			
0			
1			
2			
MENOPAUSAL STATUS			
As in meeting book			
Weight (median, range)			
Height (median, range)			
BMI			
</=30			
>30			
Mean (SD)			
HORMONE RECEPTOR STATUS			
ER-positive and/or PgR-positive (100%)			
T stage (Neo-adjuvant)			
cT1			

cT2			
cT3			
N Stage (Neo-adjuvant)			
cN0			
cN1			
cN2			
T stage (adjuvant)			
pT1			
pT2			
pT3			
N Stage			
pN0			
pN1			
pN2			
pN3			
Adj radiotherapy			
No			
Yes			
HER2			
Negative			
Positive			
Adjuvant Horome Threapy			
No			
Yes			
Adjuvant Herceptin			
No			
Yes			
Adjuvant chemotherapy			
No			
Yes			
Nodes			
Negative			
Positive			
Anti-Cancer treatment			
Yes			
No			
Cardiviscular (Yes vs No)			

Yes			
No			
Grade			
I			
II			
III			
Breast cancer subtype			
Ductal			
Lobular			
Resection margins			
Number of axillary nodes examined			
Number of positive axillary nodes			
Number of sentinel nodes examined			
Number of positive sentinel nodes			

Table 8. Baseline Hematology / Biochemistry

CTC AE 4.0 Hematology / Biochemistry table

Table 9. Baseline Non-Hematologic Adverse Event

CTC AE 4.0 Non-Hematologic Adverse Event table

Table 10. Log Rank and Cox Regression Model for IDFS Survival

Data set: All Randomized Patients				
	Univariate HR (95% CI)	Log-rank p-value	Multivariate HR(95% CI) (1)	Multivariate p-value
Treatment (Metformin vs. Placebo) (stratified)	***	***	--	--
Treatment (Metformin vs. Placebo) (un-stratified)				
BMI </= 30 versus > 30 (kg/m2)				
HER2 positive versus HER2 negative				
Chemotherapy – any versus none				

(1) Based on Cox Model with all factors included.

Table 11: IDFS by Subsets

Data set: All Randomized Patients			
	Number of patients (%)		
	Metformin N (# of event)	Placebo N (# of event)	HR (95% C.I)
ER and/or PgR			
Positive	** (**)	** (**)	** (**)
BMI			
</=30			
>30	** (**)	** (**)	** (**)
	** (**)	** (**)	** (**)
HER2			
Positive			
Negative			
Chemotherapy			
Any			
None			
Hormone therapy			
Any			
None			

Table 12: IDFS event Summary

Data set: All randomized Patients

	Number of Patients	
	Metformin N = ***	Placebo N = ***
Patients with IDFS event	***	***
Local or Regional Recurrence	**	**
Distant Recurrence	**	**
New primary (Non-BC) malignancy		
Invasive contralateral breast tumour	**	**
Death (cause 1)	**	**
Death (cause 2)	**	**
…		
Patients who were censored	*** (**)	*** (**)
Reason Censored		
No IDFS event	**	**
Withdrawal of Consent	**	**

Table 12a: BCFI event by arm

Data set: All randomized Patients

	Number of Patients	
	Metformin N = ***	Placebo N = ***
Patients with BCFI event	***	***
Local or Regional Recurrence	**	**
Distant Recurrence	**	**
Invasive ipsilateral breast tumour	**	**
DCIS ipsilateral breast tumour	**	**
Invasive contralateral breast tumour	**	**
DCIS contralateral breast tumour	**	**
Death(related to disease or treatment)		
Death (Breast Cancer)	**	**
Death (AE related to Protocol treatment)	**	**
Death (Related to non-Protocol treatment)	**	**

Table 12b: IBCFS event by arm

Data set: All randomized Patients

	Number of Patients	
	Metformin N = ***	Placebo N = ***
Patients with IBCFS event	***	***
Local or Regional Recurrence	**	**
Distant Recurrence	**	**
Invasive ipsilateral breast tumour	**	**
Invasive contralateral breast tumour	**	**
Deaths (All)		
Death (Breast Cancer)	**	**
Death (AE related to Protocol treatment)	**	**
Death (Related to non-Protocol treatment)	**	**
Death (Cardiovascular Disease)	**	**
Death (Other condition)	**	**
Death (Unknown)	**	**

Table 13. Log Rank and Cox Regression Model for Overall Survival

Data set: All Randomized Patients				
	Univariate HR (95% CI)	Log-rank p-value	Multivariate HR(95% CI) (1)	Multivariate p-value
Treatment (Metformin vs. Placebo) (stratified)	*** (*, *)	***	*** (*, *)	***
Treatment (Metformin vs. Placebo) (un-stratified)	*** (*, *)	***	*** (*, *)	***
BMI < /=30 versus > 30 (kg/m2)	*** (*, *)	***	*** (*, *)	***
HER2 positive versus HER2 negative	*** (*, *)	***	*** (*, *)	***
Chemotherapy – any versus none	*** (*, *)	***	*** (*, *)	***
Other factors…				

(1) Based on Cox Model with all factors included.

Similar tables will be reported for BCFI and distant RFS .

Table 14: OS by Subsets

Data set: All Randomized Patients			
	Number of patients (%)		
	Metformin N (# of event)	Placebo N (# of event)	HR (95% C.I)
ER and/or PgR			
Positive	** (**)	** (**)	** (*, *)
BMI			
</=30			
>/=30	** (**)	** (**)	** (**)
	** (**)	** (**)	** (**)
HER2			
Positive			
Negative			
Chemotherapy			
Any			
None			

Similar tables will be reported for BCFI and distant RFS .

Table 15: Death Summary

	Number of Patients (%) Data set: All randomized Patients		
	Metformin N =	Placebo N =	Total N =
Patients who died			
Cause of Death			
Death (Breast Cancer)			
Death (AE related to Protocol treatment)			
Death (Related to non-Protocol treatment)			
Cardiovascular Disease			
Other condition or circumstance			
Unknown			
Patients who were censored			
Reason Censored			
Still Alive			
Withdrawal of Consent			

Table 16a: Drug exposure

Data set: All treated Patients			
	Number of patients		
	Placebo N = ***	Metformin N = ***	Total N = ***
Median*	***	***	***
Min	**	**	**
Max	**	**	**

* Based on KM method, censored at IDSF event date

Table 16b: Duration of treatment by SNP status

Repeat table 20 here using the following subgroup of patients
SNP_base = 'CA' or 'CC' vs 'AA'
SNP_base = 'AA', vs SNP_base = 'CA', vs SNP_base = 'CC',

Table 16c: Drug adherence

Data set: All treated Patients			
Median (Min, Max)	Number of patients		
	Placebo N = ***	Metformin N = ***	Total N = ***
Year 1 (0-365 days)	Median (Min, Max)	***	***
Year 2	***	***	***
Year 3	***	***	***
Year 4	***	***	***
Year 5	***	***	***
Over all			

Table 16d: Drug dose modification

Data set: All treated Patients			
Dose modification	Number of patients		
	Placebo N = ***	Metformin N = ***	Total N = ***
Yes	** (**)	** (**)	** (**)
No	** (**)	** (**)	** (**)

Repeat this for each study period (years 1 to 5) and overall (0 to 5 years)

Table 17.Non-Hematology adverse events

CTC AE 4.0 Non-Hematology adverse events table.

Repeat table 17 here using following subgroup of patients
SNP_base = 'CA' or 'CC' vs 'AA'
SNP_base = 'AA', vs SNP_base = 'CA', vs SNP_base = 'CC',

Table 18. Hematology / Biochemistry (Follow-up)

CTC AE 4.0 Hematology / Biochemistry table

Table 19. Serious Adverse Events

As in meeting book table.

Table 20. Off protocol treatment

Data set: All treated Patients			
Off treatment	Number of patients		
	Metformin N = ***	Placebo N = ***	Total N = ***
Cause1	***	***	***
2	**	**	**
3			
4			

Table 21. Death on Trial within 30 days of last treatment

Data set: All Randomized Patients			
Cause of Death	Number of patients		
	Metformin N = ***	Placebo N = ***	Total N = ***
Cause1	***	***	***
2	**	**	**
3			
4			

Figure 1. KM plot for IDFS

Figure 2. KM plot for OS

Figure 3. KM plot for BCFI

Figure 4. KM plot for DRFS

Figure 5. Cumulative incidence plot for BCSM

Figure 6. KM plot for IBCFS

Figure 7a. KM plot for contralateral BC (invasive only)

Figure 7b. KM plot for contralateral BC (invasive and DCIS)