PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Exploring effects of presurgical weight loss among women with stage 0-II breast cancer: Protocol for a randomized controlled feasibility trial
AUTHORS	Tsuruta, Yuko; Rogers, Laura; Krontiras, Helen; Grizzle, William; Fruge, Andrew; Oster, Robert; Umphrey, Heidi; Jones, Lee; Azrad, Maria; Denmark-Wahnefried, Wendy

VERSION 1 - REVIEW

REVIEWER	Maurizio Montella National Cancer Institute Fondazione Pascale Napoli- Italy
REVIEW RETURNED	27-May-2016

GENERAL COMMENTS	In general, it is difficult to find the information on study details.

REVIEWER	Darren Brenner University of Calgary Alberta, Canada
REVIEW RETURNED	02-Jun-2016

GENERAL COMMENTS	This manuscript presents a study protocol for a randomized controlled feasibility trial out of the University of Alabama Birmingham. It will recruit 40 overweight or obese women, newly diagnosed with stage 0-II breast cancer, to participate in a weight loss intervention over a minimum of 3 weeks prior to surgery. In addition to assessing standard feasibility benchmarks, the authors also plan to assess the impact of pre-surgical weight loss on tumor proliferation rates, other tumor markers, serum biomarkers, gene expression, microbiome profiles and other clinical outcomes, such as quality of life (QoL). Overall, this manuscript is very well-written with good explanations of the study protocol to be used; however, there are some concerns with the hypothesis testing involved with such a small sample size, as well as the low recruitment rates already being experienced. Major Revisions - Due to low recruitment rates, the authors have opted to include pre- and peri-menopausal women. This is a concern due to the mechanistic nature of this study, as pre- and postmenopausal breast cancers operate via different mechanisms related to obesity. - Feasibility studies are designed to assess feasibility benchmarks such as recruitment potential, safety or management of the intervention, etc., not to do hypothesis testing. The authors appear

	to be very ambitious in the collection of numerous biologic samples from study participants and aim to test a large number of tumor markers, serum biomarkers, gene expression, microbiome profiles and other outcomes, such as quality of life. Despite the authors' recognition of this study being underpowered to test these in such a small study population, these biomarkers seem to be the primary outcome of the study as there is a large focus on this. There will likely be minimal real changes due to the intervention with such a small sample size and any result will likely be inconclusive. As such, the authors should acknowledge that other than the primary analyses, all other analyses exploratory, in particular the 'omics analyses where the sheer number of assessments made will increase the type 1 error rate. Minor Revisions - For the same reason of low recruitment rates, the authors have also opted to include stage I and II breast cancer patients, in addition to the original study population of ductal carcinoma in situ (DCIS) patients. There would be some concern as to the balance between the stage 0, I and II patients and if they would be matched in the randomization. - Why were only 3 days of wear selected for the accelerometer? Typical guidelines suggest 3-10 days of wear for valid data, with most recommending 7 days, especially with small sample sizes (Cain, 2013). - Under the "Power and sample size calculations" heading, it says that the initial sample size is 15 per arm, but the total sample size should be 40, thus meaning 20 per arm.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
In general, it is difficult to find the information on study details.

We appreciate the comment by the reviewer. Given that our word count is at the recommended limit, we have reviewed the manuscript and added more detail (particularly with regard to the intervention – see additions pages 12-13), and have reduced our verbiage in other areas so that the word count is still within the acceptable range.

Reviewer: 2
Overall, this manuscript is very well-written with good explanations of the study protocol to be used; however, there are some concerns with the hypothesis testing involved with such a small sample size, as well as the low recruitment rates already being experienced.

We appreciate the reviewer's comment regarding the clarity of our report. We also understand the concern regarding hypothesis testing. Indeed, this is a feasibility study, and one that is hypothesis-generating – not hypothesis testing. Therefore, we have eliminated any verbiage regarding hypotheses and have emphasized the exploratory nature of our work (see pages 5-6).

Major Revisions
- Due to low recruitment rates, the authors have opted to include pre- and peri-menopausal women. This is a concern due to the mechanistic nature of this study, as pre- and postmenopausal breast cancers operate via different mechanisms related to obesity.

We understand the reviewer's concern of combining both pre- and post-menopausal women in this feasibility study. However given that this study is not aimed primary risk, but rather disease progression from the point of diagnosis and ample evidence (see Protani ref #4) that obesity negatively affects both pre- and post-menopausal similarly, we have received approvals to go forward with this strategy.

- Feasibility studies are designed to assess feasibility benchmarks such as recruitment potential, safety or management of the intervention, etc., not to do hypothesis testing. The authors appear to be very ambitious in the collection of numerous biologic samples from study participants and aim to test a large number of tumor markers, serum biomarkers, gene expression, microbiome profiles and other outcomes, such as quality of life. Despite the authors' recognition of this study being underpowered to test these in such a small study population, these biomarkers seem to be the primary outcome of the study as there is a large focus on this. There will likely be minimal real changes due to the intervention with such a small sample size and any result will likely be inconclusive. As such, the authors should acknowledge that other than the primary analyses, all other analyses exploratory, in particular the 'omics analyses where the sheer number of assessments made will increase the type 1 error rate.

The reviewer is correct, our primary aim is to test feasibility of the weight loss intervention among women diagnosed with stage 0-II breast cancer in the pre-surgical setting. However, the innovation of trial lies in its collection of biological and other outcomes data that are collected for the purposes of hypothesis generation. We have emphasized this point within our methods and in our limitations (see pages 6 and 19)

Minor Revisions
- For the same reason of low recruitment rates, the authors have also opted to include stage I and II breast cancer patients, in addition to the original study population of ductal carcinoma in situ (DCIS) patients. There would be some concern as to the balance between the stage 0, I and II patients and if they would be matched in the randomization.

Given that this study is powered on achieving differences in weight change between study arms, we stratified on the factor that was most likely to predict that change, i.e., current BMI status. This point is clarified on page 11. We understand the concern of the reviewer however and also have stated a plan to examine effects by stage using descriptive data (see page 17).

- Why were only 3 days of wear selected for the accelerometer? Typical guidelines suggest 3-10 days of wear for valid data, with most recommending 7 days, especially with small sample sizes (Cain, 2013).

As the reviewer has indicated, we fully understand that most recommended accelerometer assessment would be 7 days. However, practically we could not afford 7 days for accelerometry to start the intervention for this pre-surgical trial since it would cause further delay of surgery and increase participant burden. We have clarified the reason behind the 3-day assessment in the revised manuscript (see page 8) and have included the citation of Cain et al. We also have specifically mentioned concerns related to accelerometry within our limitations (see page 19).

- Under the "Power and sample size calculations" heading, it says that the initial sample size is 15 per arm, but the total sample size should be 40, thus meaning 20 per arm.

The reviewer is correct and we have clarified that a sample size of 20 per arm was initially planned, though a sample size of 14-16 would still promise sufficient power to detect differences in weight loss between the study arms (see page 16).

VERSION 2 – REVIEW

REVIEWER	Darren Brenner Alberta Health Services, Canada
REVIEW RETURNED	29-Jul-2016

GENERAL COMMENTS	The authors have addressed all of the issues that were raised in the initial review of the manuscript. The current version of the manuscript greatly improved.