

## 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

<b>TITLE (PROVISIONAL)</b>	Study rationale and protocol of the BARICO study: a longitudinal, prospective, observational study to evaluate effects of weight loss on brain function and structure after bariatric surgery.
<b>AUTHORS</b>	Vreeken, Debby; Wiesmann, M; Deden, Laura; Arnoldussen, Ilse; Aarts, Esther; Kessels, Roy; Kleemann, Robert; Hazebroek, Eric; Aarts, Edo; Kiliaan, Amanda

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Yi Zhang, Professor Xidian University, China
<b>REVIEW RETURNED</b>	01-Aug-2018

<b>GENERAL COMMENTS</b>	<p>In this study protocol, the authors plan to evaluate the long-term effects of weight loss after bariatric surgery from the perspectives of brain functional and structural alterations using neuroimaging, neuropsychological, metabolic and histopathological tests. This study is worth carrying out and would be able to add new knowledge to our current understanding of Microbiota-Gut-Brain Axis.</p> <p>I have just a few comments:</p> <ol style="list-style-type: none"><li>1. With a 10 years follow-up study after bariatric surgery, I would like to suggest the authors to consider the impact of time because there were many studies showed that brain function and structure changed with age. It is necessary to recruit a well-matched control group of obese patients who did not receive surgery (Zhang et al., 2018; Liu et al., 2018, Li et al., 2018).</li><li>2. It is better to test the specific cognitive performance during the fMRI scan, based on a variety of available task experimental designs.</li><li>3. The eating behavior measures are also needed.</li></ol> <p>However, this is a very good prospective observational study protocol providing the direction of future research in obesity and bariatric surgery.</p>
-------------------------	--

<b>REVIEWER</b>	Rita Bettencourt-Silva Department of Endocrinology, Diabetes and Metabolism, Centro Hospitalar São João; Faculty of Medicine of University of Porto; Instituto de Investigação e Inovação em Saúde, University of Porto. Porto, Portugal
<b>REVIEW RETURNED</b>	26-Sep-2018

**GENERAL COMMENTS**

Page 2, Line 54: I guess there is an error ("by blood analyses"). It suggests that cognition, microbiota, etc, are measured by blood analysis.

Page 4, lines 74-75: The follow-up of 10 years is a strength of the study.

Page 5, lines 90-92: RYGB is not only a restrictive but also a malabsorptive surgical procedure. The changes in gut microbiota after RYGB in previous studies must be specified.

Page 5, lines 93-96: Several comorbidities, such diabetes, hypertension and dyslipidemia, are proven to affect cognition. Will the authors collect the data regarding the presence of these comorbidities, medication and possible remission or improvement after surgery? This should be clarified in methods section.

Page 7, lines 142-144: Please specified the 36 months of follow-up after surgery.

Page 8, lines 161-171: The aims of the study are more clear in the abstract section. In the main manuscript, the aims are more confused and the role of gut microbiota are poorly emphasized.

Page 8, line 179: Please describe the standard technique in your center and the limb lengths of the RYGB.

Page 9, lines 198-199: The lowest body mass index and the highest % excess weight loss are frequently achieved at 18-24 months after bariatric surgery. After that, there is a tendency toward weight regain. Is there any possibility to perform a MRI at 5 years of follow-up? It will add additional and important information to your study.

Page 9, lines 199-201: Some studies in this field had an important decline in the follow-up rate. Patients who experienced more satisfactory outcomes may be more likely to maintain the follow-up. I think it is a good idea try to ensure a good follow-up rate.

Page 10, lines 208-210: I suggest a study population > 40% men and > 40% women to a better equal sex distribution. A study population > 30%/30% may allow 31% men and 69% women.

Page 12, lines 270-273: 5 hours of fasting are not enough for some measurements. Please specify which vitamins and inflammatory markers will be measured.

Page 13, lines 287-304: The results of bariatric surgery is mentioned in the questionnaires section and there is other section for "physical measurements". I suggest a section with clinical data: height, weight, BMI, waist and hip circumferences, blood pressure. Formulas of %excess weight loss and % total weight loss should be present. The success of bariatric surgery in terms of weight loss should be defined. At which several time points will the authors collect anthropometric data?

Page 14, lines 326-334: Statistical analysis has to be better described. Will the authors use only linear models? Or also logistic models? What is the significance level of p value? Which statistical software will be used?

	<p>Page 16, line 374: SCFA could be replaced by short-chain fatty acids</p> <p>Figure 1: Time of medical evaluation (eg, for anthropometric measurements) should be present.</p> <p>General comments:  The manuscript must be more uniformed (4 or four, 8 or eight) and the abbreviations reviewed. Please, the first time you use an abbreviation in the text, present both the spelled-out version and the short form.  Some sentences are confused. I suggest reviewed and corrected by a native English speaker.  It is an interesting and innovative study.</p>
--	---

### VERSION 1 – AUTHOR RESPONSE

#### Answers to reviewer 1:

In this study protocol, the authors plan to evaluate the long-term effects of weight loss after bariatric surgery from the perspectives of brain functional and structural alterations using neuroimaging, neuropsychological, metabolic and histopathological tests. This study is worth carrying out and would be able to add new knowledge to our current understanding of Microbiota-Gut-Brain Axis.

I have just a few comments:

1. With a 10 years follow-up study after bariatric surgery, I would like to suggest the authors to consider the impact of time because there were many studies showed that brain function and structure changed with age. It is necessary to recruit a well-matched control group of obese patients who did not receive surgery (Zhang et al., 2018; Liu et al., 2018, Li et al., 2018).

**Reply:** *We like to thank the reviewer for this valuable suggestion. We have considered including a control group in this study. However, we are mostly interested in the effect of weight loss after surgery on brain function and structure. Therefore, patients in this study will be their own control, since 20-30% of the patients will not have sufficient weight loss after the surgery and fall back in their former, pre-surgery weight. For this reason, a control group will not be necessary. We hope that we have informed you sufficiently.*

2. It is better to test the specific cognitive performance during the fMRI scan, based on a variety of available task experimental designs.

**Reply:** *Thank you for this comment. During the fMRI scan the patient will perform a traditional Stroop test to measure the BOLD response during interference control, as well as the interference effect in response times and error rates. Due to time restriction we choose not to include more cognitive tasks in order to avoid attention deficits and discomfort of patients due to lying in the scanner for a long time, especially since we wanted to include many other advanced sequences, such as a FLAIR sequence, diffusion-weighted MRI scans and an arterial-spin labelling sequence to investigate also other important parameters such as grey and white matter integrity and cerebral blood flow.*

3. The eating behavior measures are also needed.

**Reply:** *We would like to thank the reviewer for this valuable suggestion. Patients will be asked to fill out an eating diary of two days (weekday and weekend day) on different time-points during the study. Therefore, we will have indeed an indication of the eating behaviour of the patient. Unfortunately, this was not completely clear in our manuscript, so we rephrased the respective section to clarify this point.*

To estimate the participants' food/nutrient intake and eating behaviour patients will be asked to fill out an eating diary of two days (a weekday and a weekend day).

However, this is a very good prospective observational study protocol providing the direction of future research in obesity and bariatric surgery.

**Reply:** *We like to thank the reviewer for these very nice compliments and for the useful suggestions to improve the manuscript.*

#### **Answers to reviewer 2:**

Page 2, Line 54: I guess there is an error ("by blood analyses"). It suggests that cognition, microbiota, etc, are measured by blood analysis.

**Reply:** *We would like to thank you for indicating this mistake. We have removed 'by blood analyses' in the text.*

Page 4, lines 74-75: The follow-up of 10 years is a strength of the study.

**Reply:** *Thank you for this compliment.*

Page 5, lines 90-92: RYGB is not only a restrictive but also a malabsorptive surgical procedure. The changes in gut microbiota after RYGB in previous studies must be specified.

**Reply:** *We would like to thank the reviewer for this valuable suggestion and we changed this part of the introduction. Furthermore, we have also included more information about the changes in gut microbiota that can be expected based on available literature.*

P.5

RYGB is a restrictive and malabsorptive (for micronutrients) surgical procedure; it excludes the main part of the stomach, the duodenum and the first part of the jejunum from the passage of food, leading to, among others, hormonal and gut microbiota changes.<sup>(14, 15)</sup> Gut microbiota changes after RYGB comprise increases in gut microbiota diversity, increases in relative abundance of *Actinobacteria* and *Firmicutes* phyla and decreases in relative abundance of *Bacteroidetes* phyla. However, effects in reported studies are quite inconsistent and further research is needed. <sup>(16, 17)</sup>

Page 5, lines 93-96: Several comorbidities, such diabetes, hypertension and dyslipidemia, are proven to affect cognition. Will the authors collect the data regarding the presence of these comorbidities, medication and possible remission or improvement after surgery? This should be clarified in methods section.

**Reply:** *We fully agree with this comment and following your recommendations we have added more information on the collection of data regarding the presence of the comorbidities, medication and remission after surgery.*

P.14

Furthermore, data on comorbidities like T2DM, HT and DL and associated medication will be collected before the surgery and at all time-points after surgery. Comorbidities will be defined using following criteria: for T2DM a fasting plasma glucose of  $\geq 7.0$  mmol/L and HbA1c  $\geq 48$  mmol/mol (HbA1c  $\geq 6.5\%$ ) or the use of oral antidiabetic or insulin medication; for HT the use of antihypertensive drug treatment; for DL the use of statins.

Page 7, lines 142-144: Please specified the 36 months of follow-up after surgery.

**Reply:** *Thank you for this suggestion, we specified the three years of follow-up after surgery.*

P.7

Investigators showed lasting improvements three years after surgery in the cognitive domains of attention, executive function and memory.<sup>(19)</sup>

Page 8, lines 161-171: The aims of the study are more clear in the abstract section. In the main manuscript, the aims are more confused and the role of gut microbiota are poorly emphasized.

**Reply:** *We thank the reviewer for this advice, we made the aims of the study more clear in the main text and included the role of gut microbiota.*

P.8

The primary aim of the BARICO study (**BA**riatric surgery **Rijnstate** and Radboudumc neuro**I**maging and **C**ognition in **O**besity) is to determine the long-term effect of weight loss on measures of brain function and structure after BS. The secondary aim is to provide mechanism-based rationales responsible for functional and structural decline in obese individuals. Therefore, the metabolic and inflammation status of organ biopsies will be determined together with molecular signatures via blood plasma/serum analyses. Furthermore, gut microbiota composition will be monitored over time to gain knowledge about the gut-brain axis.

This study will contribute to the development of better health campaigns, health care and preventatives to attenuate the impact of obesity. This paper describes the design and protocol of the BARICO study.

Page 8, line 179: Please describe the standard technique in your center and the limb lengths of the RYGB.

**Reply:** *We would like to thank the reviewer for this recommendation. We have added one sentence with more information on the surgical technique and limb lengths.*

P.9

A laparoscopic antecolic antegastric RYGB procedure will be performed (biliopancreatic limb of 150 cm, alimentary limb of 100 cm).

Page 9, lines 198-199: The lowest body mass index and the highest % excess weight loss are frequently achieved at 18-24 months after bariatric surgery. After that, there is a tendency toward weight regain. Is there any possibility to perform a MRI at 5 years of follow-up? It will add additional and important information to your study.

**Reply:** *We highly agree with this suggestion of the reviewer. Unfortunately, there is no funding yet to perform a MRI at 5 years of follow-up. In the upcoming years we will explore funding options to include MRI measurements at 5 and 10 years. Nevertheless, we still believe performing MRI 24 months after surgery will provide additional, important information to the performed MRI studies which have mostly a follow-up till 6 months only (Tuulari et al. 2016), because first after 24 months some patients will not have sufficient weight loss or have weight regain.*

Page 9, lines 199-201: Some studies in this field had an important decline in the follow-up rate. Patients who experienced more satisfactory outcomes may be more likely to maintain the follow-up. I think it is a good idea try to ensure a good follow-up rate.

**Reply:** *We thank the reviewer for this advice, and we are aware of this problem. We try to ensure the best follow-up rate possible via regular phone calls and feedback on the study via email and several social media sources, such as Facebook.*

Page 10, lines 208-210: I suggest a study population > 40% men and > 40% women to a better equal sex distribution. A study population > 30%/30% may allow 31% men and 69% women.

**Reply:** *Thank you for this valuable suggestion. We aim for a study population of 50% men and 50% women. However, more than 80% of our patient population in the Rijnstate Hospital consists of women, therefore 40% men is probably not achievable. This is something we are highly aware of, we assure to obtain the most optimal sex distribution possible, hopefully >40% men and >40% women. Furthermore we will correct for gender differences by including gender as factor in each statistical model.*

Page 12, lines 270-273: 5 hours of fasting are not enough for some measurements. Please specify which vitamins and inflammatory markers will be measured.

**Reply:** Thank you for your attentiveness. This was indeed incorrect, as indeed at least 8 hours of fasting is required, which we now have corrected in the text. Furthermore, we have specified the vitamins and inflammatory markers.

P.12-13

At several time points (figure 1) fasting (at least 8 hrs.) blood samples from the participants will be collected. As standard procedure classical parameters, such as several vitamins (vitamin B12, D and folic acid) and lipids (triglycerides and cholesterol) will be measured. Special interest is taken on circulating mediators of organ cross-talk, such as: cytokines, oxylipids, adipokines, hormones and inflammation markers (e.g., C-reactive protein, serum amyloid A, vascular cell adhesion molecule 1, transforming growth factor beta), as well as metabolites (derived from organs or microbiota) assessed by metabolomics, such as bile acids and bioactive (short chain) fatty acids, and other lipid species (untargeted lipidomics).

Page 13, lines 287-304: The results of bariatric surgery is mentioned in the questionnaires section and there is other section for "physical measurements". I suggest a section with clinical data: height, weight, BMI, waist and hip circumferences, blood pressure. Formulas of %excess weight loss and % total weight loss should be present. The success of bariatric surgery in terms of weight loss should be defined. At which several time points will the authors collect anthropometric data?

**Reply:** We thank the reviewer for this advice. We included more information on anthropometric measurements in the section medical evaluation.

P.14

#### **Medical evaluation**

At several time points during the study (figure 1) a medical evaluation will take place where anthropometric measurements such as: body weight, length, waist circumference and blood pressure will be quantified. BMI will be calculated as weight divided by height in meters squared. Percentage excess weight loss (%EWL) (defined as weight loss divided by preoperative excess weight, with excess weight defined as the weight above a normal BMI of 25 kg/m<sup>2</sup>) will be calculated during the time points after surgery, similar to percentage total body weight (%TBWL) (defined as weight loss divided by preoperative weight). The success of BS in terms of weight loss will be defined as a sustained weight loss larger than 50 %EWL.

Page 14, lines 326-334: Statistical analysis has to be better described. Will the authors use only linear models? Or also logistic models? What is the significance level of p value? Which statistical software will be used?

**Reply:** Thank you for this comment. We described the statistical analyses in the revised manuscript in more detail, including the significance level and the statistical software.

P.15-16

#### **Analysis of primary outcome measures**

As a primary outcome measure, baseline levels of the imaging parameters (such as MD and FA) will be compared with the results of the neuroimaging outcome 24 months after surgery (including %TBWL as a factor in the model). Next, the scores of the cognitive tests from five different time points will be analysed and compared to %TBWL. Every dependent variable will be modelled in a separate linear mixed model. %TBWL will be used as a factor. Different variables, such as: depression score, age, and gender, will be (if appropriate) included in the model. For each model, we will decide which variables to include as a factor to reduce the amount of unexplained variation. To correct for multiple outcome measures, the sequentially rejective multiple testing procedure described in Bretz *et al.* will be used.<sup>(69)</sup> Data will be analysed using SPSS (version 25 for Windows) and R (version 3.5.1 for Windows). For the cognitive tests a *p* value of <0.03 and for the imaging parameters a *p* value of <0.02 will be considered as statistically significant.

Page 16, line 374: SCFA should be replaced by short-chain fatty acids

**Reply:** In the revised version of the paper we have replaced SCFA by short-chain fatty acids.

Figure 1: Time of medical evaluation (eg, for anthropometric measurements) should be present.

**Reply:** We have added the time of medical evaluation in Figure 1.

General comments:

The manuscript must be more uniformed (4 or four, 8 or eight) and the abbreviations reviewed. Please, the first time you use an abbreviation in the text, present both the spelled-out version and the short form.

Some sentences are confused. I suggest reviewed and corrected by a native English speaker. It is an interesting and innovative study.

**Reply:** We would like to thank the reviewer for the compliment on study design. The script has been revised by a native English speaker and we have checked the script for uniformity. We believe this has significantly improved our paper. Numbers under 10 are spelt out, except for measurements with a unit or time points in the study and all abbreviations are reviewed.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Yi Zhang Center for Brain Imaging, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, China
<b>REVIEW RETURNED</b>	24-Oct-2018
<b>GENERAL COMMENTS</b>	The authors have reviewed the manuscript according to the comments.
<b>REVIEWER</b>	Rita Bettencourt-Silva Department of Endocrinology, Diabetes and Metabolism, Centro Hospitalar Universitário São João, Porto, Portugal Faculty of Medicine, University of Porto, Porto, Portugal Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal
<b>REVIEW RETURNED</b>	10-Nov-2018
<b>GENERAL COMMENTS</b>	The paper was significantly improved. I accept the manuscript but I have a doubt. Why did the authors choose a p value of <0.02 or <0.03 instead of the most used ones (0.05, 0.01, 0.005, or 0.001)? And why did they choose two different p values?

#### VERSION 2 – AUTHOR RESPONSE

**Reviewer 1:**

The authors have reviewed the manuscript according to the comments.

*Thank you for this compliment.*

**Reviewer 2:**

The paper was significantly improved.

*Thank you for this compliment and for your thorough review.*

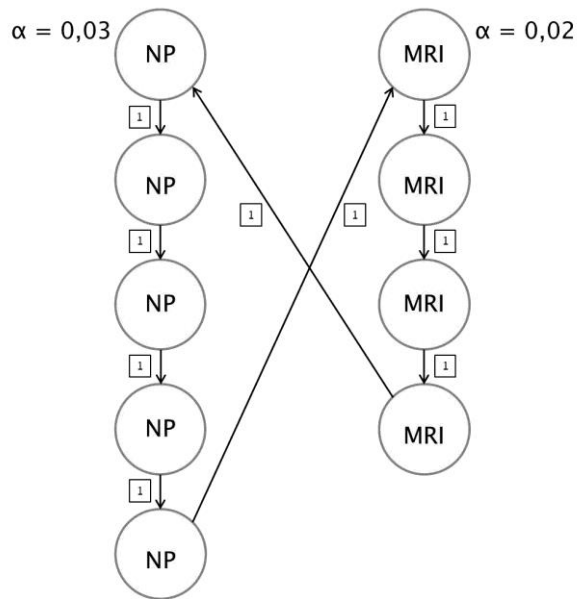
I accept the manuscript but I have a doubt. Why did the authors choose a p value of  $<0.02$  or  $<0.03$  instead of the most used ones (0.05, 0.01, 0.005, or 0.001)? And why did they choose two different p values?

*Thank you for this question. Since multiple outcome measures will be studied, we decided in consultation with a statistician to correct for this, using a sequentially rejective multiple testing procedure described in Bretz et al. (2008). Several sequentially rejective, weighted Bonferroni-based test procedures are available, which are powerful and flexible to address multiple study objectives. We did choose the most used p value of 0.05 and allocated 0.03 of this overall significance level to the neuropsychological tests and 0.02 to the MRI parameters. Since the power of the neuropsychological tests is assumed to be higher, we allocated a little more to the neuropsychological tests than to the MRI parameters. We will test the parameters sequentially based on a specific hypothesis sequence. We have included more information about the chosen significance levels in the supplementary material.*

#### **Supplementary material:**

Since multiple outcome measures will be studied, correction for this is applied using the sequentially rejective multiple testing procedure described in Bretz *et al.* (2008). As we are highly interested in both the neuropsychological tests and the MRI parameters, the MRI parameters and the neuropsychological parameters are clustered. A significance level of 0.05 is used, and an alpha level of 0.03 is allocated to the neuropsychological tests and 0.02 to the MRI parameters. The neuropsychological tests and neuroimaging tests will be tested with a multiple testing procedure (supplementary figure 1). The neuropsychological tests will initially be tested at 3/5 of the overall type I error rate (i.e. 0.03 two-sided) and neuroimaging parameters at 2/5 of it (i.e. 0.02 two-sided). Alpha will be reallocated when shown that the corresponding hypothesis is rejected. Based on the literature a specific hypothesis sequence will be tested (the sequence for the neuropsychological tests is: digit span, TAP flexibility task, story immediate/delayed recall, verbal fluency and MoCA; for the MRI parameters: DTI parameters, ASL measures, BOLD response of the Stroop test and grey and white matter volumes). Within each test separately correction for multiple testing will be included, for example for multiple brain areas analysed within a MRI parameter.





**Supplementary figure 1.** Multiple testing sequence. NP: neuropsychological tests, MRI: MRI parameters.

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Yi Zhang Xidian University
<b>REVIEW RETURNED</b>	24-Nov-2018
<b>GENERAL COMMENTS</b>	The authors has made extensive revision to the manuscript, and it can be accepted for publication.
<b>REVIEWER</b>	Rita Bettencourt-Silva 1 Department of Endocrinology, Diabetes and Metabolism, Centro Hospitalar Universitário São João, Porto, Portugal 2 Faculty of Medicine, University of Porto, Porto, Portugal 3 Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal
<b>REVIEW RETURNED</b>	23-Nov-2018
<b>GENERAL COMMENTS</b>	I accept the manuscript.