Author's response to reviews

Title: Age- and sex-specific effects on weight loss outcomes in a comparison of sleeve gastrectomy and Roux-en-Y gastric bypass: a retrospective cohort study

Authors:

Sean Manning (sean.manning@ucl.ac.uk)
Nicholas C Carter (nickcarter@nhs.net)
Andrea Pucci (andreaspucci@hotmail.com)
Alexander Jones (alex.jones@ucl.ac.uk)
Mohamed Elkalaawy (kalaawy@hotmail.com)
Wui-hang Cheung (wui-hang.cheung@hotmail.co.uk)
Borzoueh Mohammadi (Borzoueh.Mohammadi@uclh.nhs.uk)
Nicholas Finer (n.finer@ucl.ac.uk)
Alberic G Fiennes (alberic@agtwf.co.uk)
Majid Hashemi (majid.hashemi@btopenworld.com)
Andrew D Jenkinson (mradjenkinson@gmail.com)
Marco Adamo (marcadam@doctors.net.uk)
Rachel L Batterham (r.batterham@ucl.ac.uk)

Version: 4 Date: 25 June 2014

Author's response to reviews: see over
Title: Age- and sex-specific effects on weight loss outcomes in a comparison of sleeve gastrectomy and Roux-en-Y gastric bypass: a retrospective cohort study

Referee 1

We thank this reviewer for their helpful comments regarding our manuscript. For clarity, we have responded (in bold) within the original text of the comments and, we have accordingly altered our manuscript in line with the suggestions.

Reviewer's report:

Major Compulsory and Minor Essential Revisions

Lines 50-54: Please provide some information about the place where patients underwent interventions

As requested we have now added the description of the bariatric centre to the abstract (lines 50-54) and also to the study design section of the methods.

Lines 127-129: Please define hypotheses more clearly. A research hypothesis is a clear statement about an association or an effect. If you expected that age, sex and T2D could moderate the effects of procedures, you should have clarified the expectations. If you were not able to clarify any specific expectations, then you should have stated an exploratory aim. Please define your hypotheses about the moderators more clearly (clarify the expected effects) or state an exploratory aim.

We thank to reviewer for this very helpful comment. We agree that the research hypothesis and aim were not clearly defined. We have now more clearly defined our research hypothesis and stated an exploratory aim:

“We thus hypothesized that age, sex and/or T2D status could have differential effects on weight loss outcomes for each procedure. In this regard, we anticipated one of three possible scenarios: (1) that these patient factors, alone or in combination, could negatively affect the weight loss outcome in a comparable manner for both procedures (i.e. no interaction effect) (2) that these patient factors, alone or in combination, could negatively affect the weight loss outcome of one procedure but not the other (i.e. interaction effect), (3) that these patient factors, alone or in combination, could negatively affect the weight loss outcome for both procedures but to varying degrees (i.e. interaction effect). The exploratory aim of the study was to define clinical characteristics affecting weight loss, in order to optimize the recommendation of appropriate procedure for patients undergoing bariatric surgery.”
Lines 197-199: Please provide some more information about the standard post-operative timepoints (when, where and who collected weight data?). If repeated measures are equally spaced (baseline, one year and two year follow-up) and no random effect was reported, why did you use multilevel modeling? Marginal means adjusting for any covariate can be easily estimated by a repeated-measure general linear model.

We have now added information about where and who collected weight data to the section on ‘Pre- and postoperative protocol’. The details of when the weight data were collected are described in the ‘Pre- and postoperative protocol’ section. We have now also added details of the standard postoperative time points chosen for interpolation in the ‘Outcomes’ section, which were not equally spaced, thus requiring a multilevel model. As stated in the discussion, we also used a multilevel model in order to account for baseline differences in BMI between the two groups in the modelling. Baseline BMI was not a covariate in the analyses.

Lines 202-205: this sentence was uncompleted.

We thank the reviewer for noting this typo, which we have now corrected: (now line 211-213)
“Complete or partial T2D remission was defined using American Diabetes Association consensus group criteria (HbA1c <6.5%/48mmol/mol beyond one year and no active pharmacologic therapy).”

Lines 232,238,241: the first two interactions do not include time, while the third one includes time. So, why did you report results in the same manner? If the interaction do not include time, then the effects of procedure, sex and age on weight are not time-dependent because the average of repeated weight assessments is to be considered. Please clarify. Moreover, it is usual to report the model coefficients and their significance when describing the results of a multilevel analysis.

We apologise for not clarifying in the text that all interaction analyses included time. We have now clarified this. The relevant model coefficients are reported for the first interaction analysis. We have now included these for the second and third interaction analysis results also.

Level of interest: An article of importance in its field

We thank the reviewer for this assessment.

Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Referee 2

We thank this reviewer for their helpful comments regarding our manuscript. For clarity, we have responded (in bold) within the original text of the comments and, we have accordingly altered our manuscript in line with the suggestions.

Reviewer's report:
Thank you for the opportunity to review this manuscript, in which the authors try to identify preoperative factors affecting results of the bariatric procedures after 1-2 years. This study has several important limitations, many of which are not addressed sufficiently in the section on limitations in the discussion. Despite these shortcomings, this paper could be accepted for publication if the authors addressed the following comments:

1. The retrospective character of this study introduces biases in the procedure selection. Even if patients with BMI > 60 are excluded, the two groups are not entirely similar at baseline

We thank the reviewer for this comment. We have now emphasised in the limitations section of our discussion regarding bias in the procedure selection:

“... the results should be interpreted with caution due to the possibility of an unmeasured bias in the procedure selection.”

2. There are several statistics between subgroups, on the base of which authors make conclusions, or at least suppositions, but the risk of a type 2 error in this type of subgroup statistics is very high. This limitation should be addressed separately in the discussion

We agree with the reviewer that use of subgroup analyses could lead to type 2 error and have acknowledged this in the limitations section of the discussion. We also note that type 2 error occurs when an effect goes undetected due to lack of statistical power, which is not the case for the major findings in our study.

“... the subgroup analyses used in our study may have limited the power to detect important effects. For example, the lack of an effect on BMI loss observed in the T2D by procedure interaction analyses could be explained by the smaller sample size of patients with T2D in the cohort.”

3. In Fig 2, the BMI curves are presented with the SE. In this type of graph, I want the SD, which gives a much better overview of the distribution of data within the group. We know this differs very significantly between bariatric procedures
We acknowledge the reviewer’s comment regarding distribution of the data on Figure 2. We have altered the figures to demonstrate 95% confidence intervals which are generated when employing a margins plot in the statistical software program employed in the study. As stated in the figure legend, the BMI curves show the estimated marginal means at postoperative timepoints, and thus the SE and now the 95% confidence intervals relate to marginal means and not BMI data per se. Of note, the unadjusted BMI loss data at 1 and 2 years are presented with mean and SD in figure 1.

4. Your attrition rates after only 2 years are high (more than one third). This should be addressed as a serious limitation (and the fact that patients followed and not followed did not differ at baseline does not mean that they are still equal in terms of results after 2 years

We thank the reviewer for highlighting this important point. We have now emphasised this more in the limitations section of the discussion:

“…the absence of complete follow-up for approximately one-third of patients further necessitates that the results be interpreted with caution.”

5. Page 4, last line: the available studies show no difference between RYGBP and SG. OK, but you must underline that results are only short-term. No mid (3-5 years) or long (>> 5 years) results have been published. Please underline this limitations in the available literature

We thank the reviewer for this helpful comment, which we have now included in the introduction (page 4 last line).

6. Apart from the retrospective character of your study, the main limitation is your limited duration of follow-up. This limitation must be underlined clearly. You have absolutely no idea whether theses differences will remain over a longer period or no.

We thank the reviewer for making this point, which we have now acknowledged clearly in our discussion:

“… long-term weight data were not available for this study, therefore whether the findings observed persist beyond two years is unknown.”

7. The fact that several tables are available only as complementary material is unfortunate. You have to go through them, for instance, to realize that the groups are dissimilar to begin with, since the authors do not report this in their result section
We have now included the important differences in baseline criteria in the results section also. The authors felt that using Figure 1 to display the differences in important baseline characteristics for the entire cohort and for each of the subgroups analysed was the most concise way of conveying the analyses undertaken. Supplementary tables are labeled clearly in the legend of Figure 1 and throughout the text.

In other terms, many of your results are based on underpowered statistical analyses and too short a follow-up period, not to mention the fact that the two groups were not matched or randomized. The results are certainly interesting, but not sufficient to convince me that I have to change my policy in deciding which operation for which patient.

We thank the reviewer for this summary, which we have addressed in points 1, 2, 4 and 6 above. We have now toned down our conclusion regarding the anticipated impact of the manuscript regarding procedure selection, stating:

“These findings provide a strong basis for a further prospective randomized study, which could definitively address whether individualization of procedure selection, using age- and sex-specific criteria, could achieve better outcomes than standard care.”

Level of interest: An article whose findings are important to those with closely related research interests

We thank the reviewer for this assessment.

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.