Author's response to reviews

Title: Hormone Replacement Therapy is associated with Gastro-Oesophageal Reflux Disease: a retrospective cohort study

Authors:

Helen J Close (h.j.close@durham.ac.uk)
James M Mason (j.m.mason@durham.ac.uk)
Douglas W Wilson (d.w.wilson@durham.ac.uk)
A.Pali. S Hungin (a.p.s.hungin@durham.ac.uk)

Version: 2 Date: 19 December 2011

Author's response to reviews: see over
Dear Dr. Reddymasu,

Hormone Replacement Therapy is associated with Gastro-Oesophageal Reflux Disease: a retrospective cohort study

Professor Hungin and I are very grateful for the opportunity to revise and resubmit the attached manuscript. We thank the reviewers for their helpful and constructive comments. As requested, we attach a revised manuscript with all changes highlighted in yellow, and we include a point-by-point response to comments below. We would be very happy to discuss any of these further if that would be helpful.

We confirm that the manuscript has not been published previously, and is not under consideration (in whole or in part) for publication elsewhere. The manuscript, including all changes, has been approved by all authors.

Thank you for your consideration.

Yours sincerely

Helen Close
First reviewer's report:

Congratulations on a well written paper with some interesting statistical analysis. While I appreciate the authors' attempt at clarifying an important question relation to HRT and GORD, there are significant limitations in their study, which I think bear further discussion.

**Major Compulsory Revisions:**

The weaknesses primarily stem around the definition of GORD and the use of PPI prescription status:

1. GORD is often over-diagnosed and I am not convinced that this data is providing a true estimation of the incidence of GORD in the population. Not having standardized diagnostic criteria for GORD is a significant weakness, especially considering how subjective GORD symptoms are, and how much they overlap with dyspepsia and other gastric or esophageal pathology.

We acknowledge that GORD diagnosis is problematic and is a potential weakness of all retrospective studies. However, GORD diagnosis has been extensively validated and used in many GPRD studies and has been demonstrated to have more utility than self-report studies. We have added a limitation section (page 13) to the manuscript which discusses the implications of this for our study.

2. Similarly PPI use is not a reliable metric, as PPIs are not only generally over-prescribed, but are also used to treat non-GORD conditions including peptic ulcer disease, laryngopharyngeal reflux, and other forms of dyspepsia.

Similarly, we acknowledge that PPI use is not without problems. Hence, we treated GORD sufferers and PPI users as two separate groups to dilute the weaknesses inherent in each, and to allow us to fully explore the relationship between ‘proxies’ for GORD and HRT. The fact that findings were extremely similar between these proxy groups strengthens findings and is discussed further in the limitation section added to the manuscript (page 14).

3. In addition, while the statistics are impressive, they could be focused more so that it is clearer to the reader what the most relevant and important findings are. Why are significant findings only occurring in certain analyses (univariate vs adjusted etc.)? Can they comment further on this?

We are grateful for the opportunity to clarify (what we accept are) a detailed analysis process and findings. We have further explained the analysis plan and its relationship to findings firstly in the analysis section on page 7, in the results section on page 8-9, and in the discussion on page 11. This clarifies why we found different findings in different analyses which were due to the effects of matching, and other risk factors, on the strength of the association between GORD and HRT.

Second reviewer's report:

The authors Close et al have explored the association of HRT and GORD in menopausal women using validated general practice records from the UK. The authors report that oestrogen only use was associated with GORD and PPI usage.

**Strengths:**
- Well written manuscript
- Use of large well defined and validated database
- Large number of patients

**Comments:**

**Major:**
1. The proportion of missing data for important variables associated with GORD (as the authors note in the introduction) is unacceptably high (41% for smoking
and 43% for BMI). This prevented the final model to account for these factors. This should be discussed further.

We acknowledge missing data is a weakness in GPRD studies, hence the use of different analysis modelling to allow us to fully investigate the effects of different risk factors. We have added a section to the limitations on page 14 to explore this. However, we also report that sub-group analysis on patients with complete BMI and smoking data reaches the same findings as the models using the larger dataset.

2. The definition of GORD is suspect. Was this based only on the diagnostic code - gastro-oesophageal reflux? Please clarify. It should be acknowledged that this is an inadequate method to capture patients with gastro-oesophageal reflux. Lack of standardization of GORD recording as acknowledged by the authors is a major limitation.

   Please see response 2 to reviewer 1’s comments, and the reworded discussion section which addresses this.

3. Also it is unclear what classifies as GORD vs. gastro-oesophageal reflux symptoms? How did the authors account for this? This should be discussed further.

   We have clarified this, developing the methods section on page 4 in which we describe how we operationalised two different ways of defining GORD sufferers in order to strengthen our definition.

4. Limitations section should be included in the manuscript.

   This has been added to page 13-14 of the discussion.

5. How many patients underwent endoscopy in this cohort? How many had erosive esophagitis or other complications of GORD? Were those codes included?

   Endoscopy referral and findings (e.g. erosive sequelae) are not consistent with structural GORD or GORD symptoms (Montreal consensus) and were not the focus of this study which sought to understand the initial clinical presentation of GORD and its association with HRT. This is acknowledged and discussed in detail in the new limitations section on page 13.

6. Basic limitations of any retrospective cohort study such as selection bias and inability to explore temporal associations should be provided. This is especially true due to large proportion of missing data on important confounding variables.

   We agree that acknowledging the weaknesses inherent in these types of studies is important and have provided a discussion outlining the potential challenges of selection bias and temporal effects.

7. It is also unclear if PPI usage was factored into the adjusted analysis for GORD outcomes between HRT users and non users.

   Please see response 2 to reviewer 1. A discussion of this issue is provided on page 14.

Minor:
Figure 1 not available for review
We have removed the reference to this in the text.