Author's response to reviews

Title: Estimating the workload associated with symptoms-based ovarian cancer screening in primary care: an audit of electronic medical records

Authors:

Anita WW Lim (a.lim@qmul.ac.uk)
David Mesher (david.mesher@phe.gov.uk)
Peter Sasieni (p.sasieni@qmul.ac.uk)

Version: 3 Date: 31 October 2014

Author's response to reviews: see over
31 October 2014

Dear Dr Giorgi Rossi,

Re: Manuscript MS: 4462142171379730
Estimating the workload associated with symptoms-based ovarian cancer screening in primary care: an audit of electronic medical records

We thank the reviewers for their interesting feedback and comments. Please find our response below. We hope you find them to be satisfactory.

We have also corrected some minor typographical errors which have all been documented with tracked changes.

Reviewer 1

We thank Dr Tate for her comments.

a) Could the authors provide more detail on exactly what was done during the data cleaning process?

We have now added a section on data cleaning to the methods (lines 112-121 in tracked changes version). Also, further details of symptom-coding could be added as an appendix if Dr Tate or the editor feels this would be helpful.

b) It would be good if there was a reference to the large UK trial on ovarian cancer screening programme that is due to end in 2015. In particular maybe they could reference (and discuss) the interim results that were published a year or two ago. This would give the work a bit more context

We have now added a short paragraph on the UKCTOCS trial on lines 307-317 (tracked changes version).

Reviewer 2

The current NICE guidelines for ovarian cancer are based on very few primary care studies, none of which were able to report duration or persistence of symptoms. NICE recommend:

“Carry out tests in primary care...if a woman (especially if 50 or over) reports having any of the following symptoms on a persistent or frequent basis – particularly more than 12 times per month:

- persistent abdominal distension (women often refer to this as 'bloating')
- feeling full (early satiety) and/or loss of appetite
- pelvic or abdominal pain
- increased urinary urgency and/or frequency.”
Although the guidelines do suggest that particular attention should be given to women aged over 50 and to women with symptoms occurring more than 12 times per month, the wording is vague and neither are mandatory. In particular, there is no peer-reviewed evidence supporting the 12 times per month requirement. In addition, primary care clinicians do not routinely record symptom frequency or duration (recorded in 10% of consultations at the most, Hamilton et al, BJGP 2003). Therefore, while we agree that frequency and persistence are sensible to consider in the context of symptoms-based screening (and indeed in clinical assessment), we would argue that strictly adhering to a cut-off of 12 times per month is largely arbitrary and may be difficult to interpret (loss of appetite for 12 days or 12 episodes over a month). Furthermore, symptoms that prompt women to present to primary care are more likely to be frequent and persistent than those elicited by a self-reported questionnaire. As such, we believe that our approximate assessment of NICE guideline symptoms is reasonable.

Similarly, although Dr Andersen has stated that “NICE guidelines assessment should only include the group who are reporting NEW symptoms”, there is no actual stipulation in the guidelines that symptoms should be new within the past year, except for women aged 50+ with symptoms that suggest “IBS”.

**Compulsory revisions**

1) The authors mention that they do not have frequency data in passing twice in the manuscript, further comment and more discussion of the limitation of this study associated with not having frequency data would be very helpful.

Symptom frequency and duration are rarely recorded in medical records. However, we agree that this is a weakness in our dataset and may have led to overestimates of the proportion of women who would be offered targeted screening. We have now added a sentence on this in the limitations section of the discussion (lines 277-278 in tracked changes version).

2) The authors have a measure of new NICE symptoms, this is valuable having two NICE symptoms lists one of which does not reflect the new within the past year element of the guidance even inexacty is unfortunate. I would strongly encourage the authors to use just the new NICE guidelines measure they developed not the one that ignores an element of the guidance they do have data for.

Again, we agree that symptom newness is important but we think it is of interest to present data on both to give readers a feel for how much the estimates change when considering only new symptoms.

**Discresionary:**

3) Calling their moderately expanded list of symptoms the “index” symptoms is also confusing as it is not the list of symptoms used by Goff et al, for the “Symptom Index” and sounds similar. If this listing of symptoms is also well published as “Index Symptoms” the confusion is unavoidable but if this is a new nomenclature for this collection of symptoms a different name might reduce confusion among those following the field and could be desirable.

We thank Dr Andersen for raising this. We have now changed the wording to “Index 2” to match our previous publication in which “Index 2” was developed.

4) An exploration of differences between women age 45-50 and those age 50+ would also be valuable as the NICE guidelines do mention the age 50 point as important.

The proportion of consultations with NICE guideline symptoms recorded for women aged 45-49 and 50+ are detailed in the table below. Unfortunately we are unable to extrapolate
estimates of the proportion of women presenting over one year as we do not have the number of all women aged 45-49 and 50+ in each practice (age data for the practice were provided as 10 year Korner age bands; ie 45-55, 55-65 etc). However, the table below shows that estimates of the proportion with NICE guideline symptoms are broadly similar across age groups. Although a smaller proportion of women aged 45-49 were estimated to have a new NICE guideline symptom, this was based on very small numbers – only 8 women aged 50+ had a NICE guideline symptom in the week.

<table>
<thead>
<tr>
<th></th>
<th>Age 45-49 (N=365)</th>
<th>Age 50+ (N=110)</th>
<th>Age 45-74 (N=475)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (95%CI)</td>
<td>% (95%CI)</td>
<td>% (95%CI)</td>
<td></td>
</tr>
<tr>
<td>Any NICE symptom</td>
<td>6.3% (23/365)</td>
<td>7.3% (8/110)</td>
<td>6.5% (31/475)</td>
</tr>
<tr>
<td>(95%CI 4.0 to 9.3)</td>
<td>(95%CI 3.2 to 13.8)</td>
<td>(95%CI 4.5 to 9.1)</td>
<td></td>
</tr>
<tr>
<td>Proportion of NICE</td>
<td>33.3% of 23</td>
<td>49.8% of 8</td>
<td>45.0% of 31</td>
</tr>
<tr>
<td>symptoms above that</td>
<td>(95%CI 5.6 to 65.8)</td>
<td>(95%CI 23.0 to 71.9)</td>
<td>(95%CI 23.2 to 64.6)</td>
</tr>
<tr>
<td>are new*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*New within the past 1 year estimated using Kaplan Meier.

5) The addition of data of even a second single week in a different season might increase the data available allowing for both assessment of differences associated with age and better generalizability than is possible with the single week currently assessed.

We agree that this would be advantageous however, primary care medical record data are time consuming to gain permission to access and to extract, clean and analyse. Moreover, the aim of the audit was to get a rough snapshot of the primary care workload associated with symptoms-based ovarian screening in the UK. We feel that the inclusion of four different GP practices in different areas and on different weeks has provided sufficiently informative results which meet the study aims.

6) If it were possible to get data on the % of women reporting a NICE symptom who report it to be frequent (more than 12 times a month) for adjustment of the estimates would enhance the point in your paper’s discussion section which does mention that the NICE symptoms would trigger screening only if frequent would also help considerably.

Unfortunately these data are not recorded in GP medical records. Although we have access to symptom frequency data collected via self-completed questionnaire and interview from a previous study, we do not think this would be useful for two reasons:

1) symptom reporting varies considerably depending on the method of elicitation (symptom prevalence is much higher in self-complete questionnaires than on medical records), therefore adjustments using symptom data collected using questionnaire or interview are unlikely to be accurate/relevant

2) frequency data on retrospectively reported symptoms is difficult to interpret as it is often unclear for which period the frequency was rated for: when the symptom started, at the time it was worst, an average over the entire symptom duration, now, etc.

A much more minor point: Scare quotes around “Targeted” on line 132 seem unnecessary. The targeting might be poor or not and how good it is, is a topic of discussion in this paper but it is targeting.

This was a mistake. When referring to targeted screening in our manuscript we had intended to put quotes around the word ‘screening’. The reason for this is because some would say
that offering a test to a patient with symptoms is not screening. We have now been through the entire manuscript and corrected this (lines 58, 130, 186 of tracked changes version).

Yours sincerely,

Dr Anita Lim