Manuscript ID: BMJ.2015.028305

Point-by-Point Responses to the Editorial Board and Reviewers' Comments

- **Response to the Comments from the Editorial Board**

1. ***First and foremost, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below. The paper is not entirely clear on why this should be considered overdiagnosis, and to what extent. The claims you currently submit can't easily be linked to the data presented in the paper. Please lay out the arguments more clearly, in the point-by-point response to reviewers and editorial comments, as well as in the paper.***

.

Thank you very much for favorable consideration given to our manuscript. We have revised the manuscript as indicated in our responses to accommodate the suggestions and comments. Additional changes were incorporated in the revised manuscript to correct and clarify the meaning of the context. All the changes are highlighted in yellow.

1. To better address the issue of "overdiagnosis", we included supplementary Figure 1 in the main body of revised manuscript, which showed rapidly rising incidence of thyroid cancer against the steady state mortality rate over time in Korea. (Page 6, Paragraph 2 , Lines 18-19)

2. We also re-iterated this issue by summarizing the results of our study in the Discussion

Page 14, Paragraph 2 Line 14-22:
Our study showed that the great majority of the recent increases in the incidence of thyroid cancer in Korea was due to more detection of small-sized (<20 mm) tumours, which accounted for 94.4% of the overall increase in the estimated thyroid cancer incidence between 1999 and 2008. **By the SEER summary stage, 97.1% of the increase in the estimated thyroid cancer incidence was due to increased detection of regional stage tumours (61.6%) and localized stage tumours (35.5%), for which 5-year relative survival rates were 100.1% and 100.4%, respectively in Korea.[21]** Obviously, a large portion of this increase was attributed to the widespread practice of thyroid cancer screening with ultrasonography, which started around the turn of the century in Korea. [8, 22]

Page 15, Paragraph 3 Line 20- Page 16, Paragraph 1 Line 3:
Taken together with the accumulating data that showed the thyroid cancer mortality rates remained stable for decades despite of the rapid increase in its incidence rate (Figure 1), [8,24] our findings provide further supporting evidence for the overdiagnosis as a cause of thyroid cancer epidemic in Korea. **Otherwise, there is no good explanation for the findings that about 60% of the total increase in thyroid cancer incidence between 1999 and 2008 occurred in a short time span of 3 years between 2005 and 2008, and the rising trend continued even thereafter as shown in Figure 1.**

2. ***Follow-up seems too short to know conclusively whether these small tumours constitute overdiagnosis. This may be an unavoidable limitation, but worth noting and discussing in the paper.***

To accommodate your suggestion, we added a new sentence in the Discussion Section as follows;

"Nevertheless, because of relatively short duration of follow-up, we could not secure the long-term survival outcome data, which is the inherent limitation for the study of thyroid cancer." (Page 16, lines 22 – Page17, Line1-2)

3. Several editors commented that the paper may not be adding enough for a general journal to prioritise. Perhaps you could make the novel points clearer in the text, however we are of course not asking you to go overboard. We considered one of the reviewers' suggestion - to do more on spread to the lymph nodes - a good point on how to improve the paper's value in terms of novelty.

Regarding the issue of novelty, we revised the "what this study adds" Section as follows (Page 5, Line 15-18);

- We rephrased the bullet point #3 to add new findings of regional SEER stage subgroup analysis, as follows,

  " Thyroid cancer screening can detect notably small-sized tumours, but also clinically indolent asymptomatic tumours with local extension and lymph node involvement."

A new bullet point was added to emphasize the merit of our study regarding the issue of overdiagnosis

- Our study provides clear evidence that the increase in the incidence of thyroid cancer in Korea was mainly due to overdiagnosis.

Regarding the issue of more in-depth analysis of "spread to the lymph nodes", we added **a new Table (Table 3)**, a new sub-section, named "**Subgroup analysis of regional SEER summary stage tumours" (Pages 13-14) and also a new figure, Figure 3C.** Inclusion of this data in the revised manuscript further strengthened our position regarding overdiagnosis as a cause of thyroid cancer epidemic in Korea.

First, it demonstrates that there was no difference in lymph node involvement by the route of detection. Second, there were clinically indolent asymptomatic regional SEER stage tumours with local extension and lymph node involvement detected by screening, which has been shown to have no survival advantage over the local SEER stage tumours [21].

Response to the Comments from Reviewer #1.

1.1. Full text Table I: please provide p-values for differences between groups.

.

As recommended, revised the Table 1 to add p-values for differences between the groups in the last column. We moved the position of the column for total number to accommodate this change.

1.2. Table 1: histologic types--does the decrease in percentage of follicular carcinoma from 1999 (7%) to 2008 (1.8%) represent a true decrease or does it represent more follicular lesions being called Follicular Variant of Papillary Carcinoma? This needs a comment.

Response)

Histological types of thyroid cancer were classified according to the *International Classification of Diseases for Oncology,* 3rd edition (ICD-O-3) as papillary carcinoma (ICD-O-3 codes 8050, 8260, 8340–8344, 8350, and 8450–8460), medullary carcinoma (ICD-O-3 codes 8345 and 8510–8513), follicular carcinoma (ICD-O-3 codes 8290 and 8330–8335), anaplastic carcinoma (ICD-O-3 codes 8020–8035), others (ICD-O-3 codes 8000–8005, 8337, 8346, and 8347).[ref 20]

In our study, follicular variant of papillary carcinoma (8340) was classified as papillary type carcinoma. The decrease in percentage of follicular carcinoma demonstrates a true decrease in proportion due to the increase in papillary carcinoma cases in 2005 and 2008.

1.3. I think that Supplementary Figure 1 should be included in the text.

As recommended, we included supplementary Figure 1 in the main body of revised manuscript. (Page 6, Paragraph 2, Lines 18-19) During this process, we realized that previous version of the figure included the incidence rate for women in Korea. In the revised manuscript, we corrected this error by including the incidence data for all population, both women and men.

1.4. This study emphasizes the need to try and differentiate between indolent tumors and the more aggressive tumors; otherwise there will be many patients undergoing unnecessary surgeries for tumors that would never have caused problems.
One approach is through the use of molecular markers in the FNA specimens, while another is the active surveillance of the small papillary carcinomas as described by Ito and colleagues (Ito Y, et al. Thyroid 2003;13:381-7) as well as others. A brief discussion of this should be added to the Discussion section to allow readers to think about ways to handle the overdiagnosis issue without sending all patients to surgery..

Although we agree the reviewer's comment on the importance of distinguishing the indolent tumors from more aggressive tumors, our study focused more on the issue of overdiagnosis of thyroid cancer by screening. Distinction between indolent tumors and more aggressive tumors is the issue of clinical management after detection of thyroid nodules, which is beyond the scope of our study.

However, we added a few lines to address the issue of clinical management of small-sized papillary thyroid cancers as suggested by the reviewer, as shown below in the Discussion Section (Page 18, Paragraph 2 , Lines 11-14)

"Generally, only nodules >1 cm were recommended for further evaluation, since they have a greater potential to be clinically significant cancers [23]. If there is no evidence of clinical progression of tumour, some investigators recommended clinical observation for small-sized papillary thyroid cancers because they do not usually become more aggressive form.[29, 30]"

1.5. Minor corrections:

Pg 12, line 14--remove the first "tumor" from the sentence.

Pg 13, line 52--The sentence starts with "Moreover" but there is no rest of the sentence.

Pg 16, last sentence--should connect with prior sentence and connect with a comma.

We appreciate the reviewer for thorough review and pointing out our oversight in fine details. We removed the first "tumor" from the sentence and the word "moreover".
We also removed "these increases. while" .

Response to the Comments from Reviewer #2.

2.1. The work adds to the growing body of literature on overdiagnosis. It brings out in more detail the terrible problem of overdiagnosis of thyroid cancer that is occurring in Korea and was first brought to light for people in United States by the New England Journal of Medicine article last year. The paper focuses quite a bit on the change in the size of detected tumors. This is very similar to the approach that we took the last time we looked at the SEER data when we published our paper (Davies & Welch) in JAMA Otolaryngology in 2014. I think this group is missing a great opportunity by focusing on tumor size.

I think that this group has very interesting and unsettling data that they could highlight better in the dramatically increasing incidence of detected regional spread of thyroid cancer to the lymph nodes.

Presuming that the regional spread of disease is also a product of overdiagnosis, these results call into question our current thinking about how we might best manage thyroid cancer overdiagnosis. In the thyroid cancer field, the discussions right now are about whether we can monitor cancers of one cm, or 1.5 cm.

If it turns out we also need to be thinking about whether people with regional spread should be observed, that is a big shift in thinking. Focusing on the importance of the fact that we are detecting regional spread of disease without a change in mortality would really be an important new piece of data.

Presuming that the authors agree with my suggestion of how they should refocus this paper to point out the dramatic increase in the detection of regional involvement that still does not increase mortality, I think this paper would be very important. If they feel it is important to maintain a focus on tumor size, it will still out a lot, but it will not be as impactful.

We totally agreed with the reviewer that the abrupt and rapid increase in thyroid cancer was due to the overdiagnosis. In fact, we were quite motivated by the JAMA Otolaryngology paper [24]. As the reviewer pointed out we could have missed a great opportunity if we focused only on tumor size. Following the recommendation of the reviewer, we further analyzed the subgroup of tumors in the regional SEER stage. We found that there was dramatic increase in the detection of regional lymph node involvement by screening, which has been shown to have no survival advantage over the local SEER stage tumours in Korea.

In the revised manuscript, we added a new Table (Table 3) and a new sub-section, named "Subgroup analysis of regional SEER summary stage tumours" (Pages 12-14) and also a new figure, Figure 3C. Also, we added a line in the Discussion which reads as follows.

"By the SEER summary stage, 97.1% of the increase in the estimated thyroid cancer incidence was due to increased detection of regional stage tumours (61.6%) and localized stage tumours (35.5%), for which 5-year relative survival rates were 100.1% and 100.4%, respectively in Korea.[21]" (page 14, paragraph 2, lines 17-20)

2.2. The research question is clearly defined. There are some issues with the presentation of results that need to be cleared up. The methods lacked some key details that would be helpful to most readers. First, the staging system is the AJCC staging system, not TNM. It should be called the AJCC sixth edition, not the TNM sixth stage.

As pointed out, we changed the TNM 6th edition to AJCC 6th edition (Table 1; page.8, Line 16; page.9, Line 18; page.10, Line 4) .

2.3. The use of the SEER summary stage variable requires explanation in the form of a clear definition of local, regional, and distant. Only those of us who work closely with SEER data know what it means, and in fact the definition of local, regional and distant varies by tumor site. Therefore, it is really important to indicate what exactly constitutes 'regional spread' in the seer summary data variable for thyroid cancer.

As pointed, it is quite important to clarify what the SEER summary stage means in our study. We added a sentence in the Methods Section (page.8, paragraph 3, Line 17-21) as follows.

The SEER summary stage grouped thyroid cancers in 3 major categories, localized, regional, and distant and the regional stage includes 1) regional by direct extension only, 2) regional lymph nodes involved only, and 3)regional by BOTH direct extension AND regional lymph node involved. [17]

2.4. ***Specifically, I would guess that regional spread means involvement of lymph nodes. If this is the case, there is potentially a problem with table one.***

Regional lymph node involvement in table 1 does not appear to have change substantially between 1999 and 2008, but the seer summary stage for regional steadily increased from 1999 to 2008. This suggests that either the definition is not what I thought it was for thyroid cancer, or there's an error in the table. The mismatch between the Seer summary stage numbers for table 1 and table 2 also feels unsettling. The numbers are being expressed differently one is case numbers and the other is age standardized incidence rates, but somehow I feel like they should probably match up better than they do.

As the reviewer pointed out, Table 1 showed the number of cases and percentage while Table 2 showed the age-standardized estimates of thyroid cancer incidence for a given year. However, it should be noted that different sampling proportions were applied for each study year (33% in 1999, 22% in 2005, and 11% in 2008) because the number of cases diagnosed in 1999 and 2005 was smaller than that in 2008, as clearly stated in Methods Section (Page 8, paragraph 2, Line 13). Obviously, this difference in sampling proportions was reflected in the estimation of the incidence in different years, which appeared to be *"**mismatch between the SEER summary stage numbers for table 1 and table 2"***. In addition, there are additional cases grouped in regional SEER stage due to minimal extrathyoidal extension even without lymph node involvement, as shown in the newly added Table 3.

2.5. The results are somewhat confusing in places because the data that are presented visually do not always match data that are reported in the results text. For example, the authors refer on page 11 to a supplementary table 2. This is said to be data on regional lymph node involvement over time by tumor size. There is no such table in the data set, and even looking to see if something was mislabeled, I could not find any display of these data. This is too bad, because I'm very interested in seeing these data.

We are sorry, it was our fault not uploading the supplementary tables. We've now added the supplementary tables along with the revised manuscript.

2.6. Table 1 is clear and useful. Table 2 also provides useful data. Figure 1 is helpful, but might benefit from a clear divider to separate the screen detected plots from the clinical detected and total. I personally find figures 2A and 2B to be very hard to interpret. It needs a label on the Y axis, but also I'm not sure what I'm supposed to get from this figure. It may need a different method of illustration.

As suggested, we added a label on the Y axis in Figure 3. In addition, Figures 2A and 2B (which are now changed to 3A and 3B in revised manuscript) are revised to make it more friendly to the readers.

2.7. Supplementary figure 1 should be part of the primary data in my opinion. Even though it is a familiar figure to many, it is still very striking and helps to place the problem of overdiagnosis in perspective.

As recommended, we included supplementary Figure 1 in the main body of revised manuscript. (Page 6, Paragraph 2 , Lines 18-19) During this process, we realized that previous version of the figure included the incidence rate for women in Korea. In the revised manuscript, we corrected this error by including the incidence data for all population, both women and men.

Response to the Comments from Reviewer #3.

3.1. Further clarification for what constitutes "screen-detected" versus "clinically detected" cases would be helpful. For the former, the assumption is that these are cases discovered by screening ultrasound, and not other forms of screening such as physical exam or radiologic studies done for other reasons (incidental findings)? For the latter, more details regarding what tumor-related symptoms led to clinical detection would be informative (pain?, detection of a painless lump?, etc…).

We fully agree with the reviewer that ***clarification for what constitutes "screen-detected" versus "clinically detected" cases*** is very critical point. We used a pre-designed checklist to sort out the route of tumour detection as described in the medical records, including the circumstances why the patients sought medical attention. All the cases who came for a screening purpose without any thyroid-related symptom were grouped as screen detection and any cases who had records of any thyroid-related symptoms, known history of hypothyroidism or any non-specific complaint such as easy fatigability, axillary discomfort were grouped as clinical detection. Because of the retrospective nature of the study, unfortunately, no information was available in 23.4% of the cases, which was grouped as unspecified, as shown in Table 1. We addressed this issue in the Discussion Section (Page 16, Paragraph 3) by stating that "Our data may have a misclassification bias regarding the routes of tumour detection, which may cause either underestimation or overestimation of incidence rate in specific subgroups." However, we think it is unlikely to introduce any significant bias in estimation of overall thyroid cancer incidence due to misclassification of the detection route.

3.2. I certainly agree with the author's central premise that increased screening is primarily responsible for the overdiagnosis of clinically insignificant thyroid cancers, and this is the predominant cause for the rise of thyroid cancer incidence in South Korea. The three publications highlighted by the authors (refs 11-13) as challenges to this premise, however, were not analyses performed in the Korean population and did not completely discount the primary contribution of overdiagnosis, but suggested that other factors may be minor contributors given the observation of small increases in more advanced disease. The authors do also make note of minor increases in larger thyroid cancers, extrathyroidal invasion, and to a lesser extent distant metastases (in the screening group of Table 2) in their own data set. Is it possible that in addition to overdiagnosis there is also a small, minor increase in more advanced disease as well that may speak to other contributors?

Yes, it is possible as pointed by the reviewer that in addition to overdiagnosis there is also a small, minor increase in more advanced disease as well. However, given the magnitude of contribution of screen-detected thyroid cancer cases as shown by our study and also by other investigators, it remains to be seen whether there is real contribution of other yet unidentified factors. In fact, there is no better explanation than overdiagnosis for the findings that about 60% of the total increase in thyroid cancer incidence between 1999 and 2008 occurred in a short time span of 3 years between 2005 and 2008, and the rising trend continued even thereafter, as shown in Figure 1 of the revised manuscript.

3.3. The results regarding regional lymph node involvement needs greater clarification. The text

refers to a Supplementary Table 2, but none was included in the manuscript that I reviewed.

We are sorry, it was our fault not uploading the supplementary tables. We've now added the supplementary tables along with the revised manuscript.

3.4. Criteria for what constitutes lymph node involvement would need to be included. While there is an increase in nodal involvement noted over time, I suspect this is related to greater detection of small regional nodes that were of marginal clinical significance?

As the reviewer point out, it is important to further investigate the clinical significance of lymph node involvement by its extent in the involved lymph node. However, this study did not examine this aspect of lymph node involvement. But we cited the survival outcome data (Ref # 21, newly added in the revised manuscript) which showed no significant difference in relative 5-year relative survival rates, 100.1% for regional stage tumours and 100.4% for localized stage tumours.

3.5. Would it be possible to break down the regional stage tumor category (according to the SEER criteria) to disease that qualified because of lymph node involvement only versus those that did because of direct extension? Likely, the increase in regional stage over time is the result of clinically insignificant nodal involvement, but clearly delineating this and establishing that an increase in extensive, invasive primaries did not contribute would be important to support the central premise of the paper.

As recommended, we reanalyzed our data and added a new Table (Table 3), entitled " Estimated age-standardized incidence rate* of thyroid cancer with regional stage by the degree of extension and lymph node involvement according to the routes of tumour detection, 1999-2008"

We also added a new sub-section, named "Subgroup analysis of regional SEER summary stage tumours" (Pages 12-14) and also a new figure, Figure 3C. Inclusion of this data in the revised manuscript further strengthened our position regarding overdiagnosis as a cause of thyroid cancer epidemic in Korea.

As indicated in our response #3.4, our results are consistent with the reviewer's prediction that ***the increase in regional stage over time is the result of clinically insignificant nodal involvement.*** These findings further support the central theme, overdiagnosis is the cause of thyroid cancer epidemic in Korea, and also in other countries.