Author's response to reviews

Title: A network meta-analysis of therapeutic outcomes after new image technology-assisted transurethral resection for non-muscle invasive bladder cancer: 5-aminolaevulinic acid fluorescence vs hexylaminolevulinate fluorescence vs narrow band imaging

Authors:

Joo Yong Lee (joouro@yuhs.ac)
Kang Su Cho (kscho99@yuhs.ac)
Dong Hyuk Kang (doudong14@gmail.com)
Hae Do Jung (urooru@yuhs.ac)
Jong Kyou Kwon (jkstorm@naver.com)
Cheol Kyu Oh (ckohuro@gmail.com)
Won Sik Ham (uroham@yuhs.ac)
Young Deuk Choi (youngd74@yuhs.ac)

Version: 3
Date: 6 June 2015

Author's response to reviews:

Dear Editors of BMC Cancer:

Thank you for your thorough review of our manuscript (MS: 1490642980162896) entitled “A network meta-analysis of therapeutic outcomes after new image technology-assisted transurethral resection for non-muscle invasive bladder cancer: 5-aminolaevulinic acid fluorescence vs hexylaminolevulinate fluorescence vs narrow band imaging.” We are grateful for the chance to submit a revised version of our manuscript for publication in BMC Cancer. Our manuscript has been carefully revised according to the reviewers’ comments. Please find our responses to the reviewers’ comments beginning on the next page.

We hope that our revised paper is now deemed acceptable for publication, and we look forward to receiving your final decision.

Sincerely,
Young Deuk Choi, M.D., Ph.D.
Department of Urology, Severance Hospital, Urological Science Institute, Yonsei University College of Medicine,
50-1 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea
Tel: +82-2-2228-2317 / Fax: +82-2-312-2538 / E-mail: youngd74@yuhs.ac

Referee 1:
The authors apply network meta-analysis to make direct and indirect
comparisons of photodynamic and narrow band imaging modalities in NMIBC. The study question has value since these modalities have not been previously compared head to head. There are few major limitations that limit the impact of the paper and interpretation of results:

MAJOR

Comment 1: Since this is a network analysis that compares interventions that have not been previously compared directly, a thorough network geometry map/figure is needed to understand the interactions between the various nodes being compared. Figure 2 in its current format is not useful and should be replaced with a figure showing the number of studies that are being evaluated for each comparison. Refer to Mills et al BMJ 2013;346:f2914 doi: 10.1136/bmj.f2914 for further details.

Answer 1: Thank you for the comment. The network plot should include the interactions between the various nodes being compared according to your comment. Thus, we changed Figure 2 to the figure included in the article by Mills et al. (doi: 10.1136/bmj.f2914)

Comment 2: Were there any studies that compared more than one new imaging modality with WLC? Table 7 suggests that all studies compared one agent with WLC, but the results (line 107, page 10) suggest that some studies may have had more than 2 arms.

Answer 2: Thank you for the comment. We corrected the sentence as follows to eliminate any confusion:

# These studies covered therapeutic outcomes of TUR assisted by three different types of PDD or NBI versus WLC (Fig. 2).

Comment 3: Table 1: should include OR/HR and p-values from these studies. I don’t think the age column in this table is needed. Also NIT is incorrectly labelled as "now" image technology in the legend. If "quality assessment" in this table means cochrane risk of bias, then this should be clearly shown.

Answer 3: Thank you for the comment. The OR/HR from each study was not described in every paper; thus, we cannot add OR/HR in Table 1. In my opinion, these studies had a relatively short-term outcome between two-arm interventions. So, in each study, the authors could not analyze outcomes using a Cox proportional hazards regression model. In half of the studies, recurrence-free rates were determined using Kaplan-Meier curves (KMCs); however, extraction of HR/OR from KMCs cannot guarantee accurate HR/OR. Thus, again, we cannot add OR/HR in Table 1. Age was deleted according to your comment. NIT was corrected as “new image technology.” Quality assessment had a footnote as follows:

# aQuality assessment was based on Cochrane’s risk of bias as a quality assessment tool for RCTs. If four or more domains are deemed “unclear” or “no,” the study was classified as having a high risk of bias. If two or three domains were deemed “unclear” or “no,” the study was classified as having a moderate
risk of bias

Comment 4: Please combine Figure 8 and Table 2 into a single figure. These refer to the same analysis and should be shown together. Figure 8 on its own is not sufficiently informative

Answer 4: Thank you for the comment. Fig. 8 and Table 2 were combined.

Comment 5: The figure legends in general for ALL the figures are not detailed enough. The figures are not well labeled and lack legends. Specifically: Fig2 already discussed above

Fig 4: needs legend, what do symbols mean?

Fig4 is sufficient and Fig3 unnecessary

Fig 5: funnel plots need to be explained to lay reader, what are the axes, what do these figures mean?

Fig 9: explain this figure better in legend and in results section. How to interpret this for general reader.

Answer 5: Thank you for the comment. The symbols of Fig. 4 indicated low risk of bias (Green), high risk of bias (Red), and unclear risk of bias (yellow), according to guidelines on Cochrane’s risk of bias. Legends were added in Fig 4. According to your comment, Fig 3 was deleted. A funnel plot is a graph designed to check for the existence of publication bias, and funnel plots are commonly used in systematic reviews and meta-analyses. In the absence of publication bias, it assumes that the largest studies will be plotted near the average, and smaller studies will be spread evenly on both sides of the average, creating a roughly funnel-shaped distribution. Deviation from this shape can indicate publication bias. We added an explanation of the funnel plot. In Fig 9, the rank plot demonstrates the posterior probability of the network meta-analysis. Mostly, 5-ALA was ranked first and NBI was ranked second. These findings suggest that 5-ALA may be more effective than NBI and HAL.

# Fig. 3. Risk of bias summary. Review authors’ judgments for each risk of bias item for each included study. Green, low risk of bias; Red, high risk of bias; and Yellow, unclear of risk of bias.

# Publication bias was examined using funnel plots. In the absence of publication bias, this method assumes that the largest studies will be plotted near the average and that smaller studies will be spread evenly on both sides of the average, creating a roughly funnel-shaped distribution. Deviation from this shape can indicate publication bias. Quality assessment and investigation of publication bias were carried out using Review Manager 5 (RevMan 5.2.3, Cochrane Collaboration, Oxford, UK).

Comment 6: The heterogeneity assessment is shown in very mathematical terms that mean very little to the average reader (page 11, lines 121-130). The reader needs a better idea of the heterogeneity of the RCTs being compared. For example, was there heterogeneity amongst the RCTs comparing ALA vs WLC or HAL vs WLC etc. This needs to be clearly reported so that the reader can gauge
the certainty of the findings of the meta-analysis.

Answer 6: Thank you for the comment. In Figs. 6 and 7, heterogeneity test revealed little heterogeneity in our study. In the Methods section, we added a description of heterogeneity test as follows:

# An I² # 50% was considered to represent substantial heterogeneity. For the Q statistic, heterogeneity was deemed to be significant for p less than 0.10 [6]. If there was evidence of heterogeneity, the data were analyzed using a random-effects model. Studies in which positive results were confirmed were assessed with a pooled specificity with 95% CIs.

Comment 7: Discussion: Importantly, the finding that ALA is superior to HAL needs more discussion. Is this a statistical anomaly due to smaller sample size? The individual trials show similar OR for HAL vs WLC and ALA vs WLC, so how is it possible that the indirect comparisons show difference of HAL vs ALA?

Answer 7: Thank you for the comment. In pair-wise meta-analyses, 5-ALA vs WLC showed OR 0.34 95% CI 0.22-0.51, however, HAL vs WLC showed OR 0.58 95% CI 0.45-0.74. In OR, 5-ALA is lower than HAL in conventional meta-analysis. In network meta-analysis, OR was 0.48 which was similar to conventional meta-analysis, however, 95% CI was longer than conventional meta-analysis. Long 95% CI of network meta-analysis was calculated by indirect comparison based on Bayesian networking. We added this to the Discussion as follows:

# In pair-wise meta-analyses, 5-ALA versus WLC showed an OR of 0.34 and 95% CI of 0.22-0.51; meanwhile, HAL versus WLC showed an OR of 0.58 95% and a CI of 0.45-0.74. In regards to ORs, that for 5-ALA was lower than that for HAL in conventional meta-analysis. In network meta-analysis, the OR was 0.48, which was similar to that in conventional meta-analysis; however, the 95% CI was longer than that in conventional meta-analysis. The longer 95% CI for the network meta-analysis was calculated by indirect comparison based on Bayesian networking.

MINOR
- page 9, line 92 : typographic error
- page 10, line 101: Do they mean 410 articles not 41
- The statement on page 12 line 153 conflicts with the statement on page 12 line 149
- page 15, line 212: "NBI guided TUR is preferable".......compared to what and in what situation?
- Fig 2 legend: the last sentence starting "in four studies..." does not make sense.

Answer 8: On page 9, the typographic error was corrected. On page 10, 410 was changed to 41. On page 12, the two sentences were changed according to your comment. On page 15, it was revised as “NBI-guided TUR is preferable to PDD”. The legend for Fig. 2 was also changed.
# Model fit was appraised by computing and comparing estimates for deviance and information criterion.

# The database search found 41 articles covering 398 studies for potential inclusion in the meta-analysis.

# Cancers resected using 5-ALA-based PDD occupied the highest rank in the rank probability test for recurrence rate, followed by those resected using NBI (Fig. 9A). NBI-assisted TUR was ranked highest in the rank probability test for progression-free rate, followed by TUR using HAL-based PDD (Fig 9B).

# NBI-guided TUR is preferable to PDD because the specificity of PDD significantly decreases in patients who have undergone a recent instillation. Only four studies of two arms of TUR with NBI and WLC have been published.