Author’s response to reviews

Title: Do we know enough about the effect of low dose computed tomography screening for lung cancer on survival to act? A systematic review, meta-analysis and network meta-analysis of randomised controlled trials.

Authors:

Huiqin Yang (H.Yang@exeter.ac.uk)
Jo Varley-Campbell (jo.varley-campbell@ucl.ac.uk)
Helen Coelho (H.Coelho@exeter.ac.uk)
Linda Long (L.Long@exeter.ac.uk)
Sophie Robinson (S.R.Robinson@exeter.ac.uk)
Tristan Snowsill (T.M.Snowsill@exeter.ac.uk)
Ed Griffin (E.A.Griffin@exeter.ac.uk)
Jaime Peters (J.Peters@exeter.ac.uk)
Chris Hyde (C.J.Hyde@exeter.ac.uk)

Version: 1 Date: 03 Jul 2019

Author’s response to reviews:

DAPR-D-18-00034

Do we know enough about the effect of low dose computed tomography screening for lung cancer on survival to act? A systematic review, meta-analysis and network meta-analysis of randomised controlled trials.

Huiqin Yang; Jo Varley-Campbell; Helen Coelho; Linda Long; Sophie Robinson; Tristan Snowsill; Ed Griffin; Chris Hyde Diagnostic and Prognostic Research

Response to review comments

Reviewer #1:

Comment 1.
The authors tried to answer the question "Do we know enough about the effect of low dose computed tomography screening for lung cancer on survival to act?", however, fell well short.

Response 1.

We have focused on examining the evidence on mortality because we felt there were some complexities to this which had not been sufficiently debated. We remain content that although mortality data is not sufficient to proceed with implementation of screening for lung cancer with LDCT, it is necessary, so limitations with the mortality data alone should lead to caution with implementation. We are thus content that we have addressed the question posed in the title.

Change 1.

No change

Comment 2.

This review provides no additional information as compared to a recent systematic review by Ali et al 2016 on "Screening for lung cancer: A systematic review and meta-analysis".

Response 2.

Whilst Ali et al 2016 is an excellent systematic review the research we report advances knowledge in several ways. a) It updates the review from March 2015 to June 2017. As a corollary we provide more up-to-date outcome data for DLCST b) We provide more information on the trajectory of the trials in progress, important given that part of our conclusions point to the evolving nature of the evidence on mortality of LDCT screening for lung cancer c) We extend the standard RCT quality assessment to include a detailed examination of baseline equivalence with important concerns noted about the MILD RCT d) We use network meta-analysis to formally explore the influence of the comparator on estimates of effectiveness of LDCT e) We explore the heterogeneity in the results noting that this has been ignored by other reviewers including Ali et al. f) We offer a less optimistic estimate of the effect of LDCT on mortality based on all the trial evidence, rather than focusing on NLST alone which we believe is methodologically incorrect

Change 2.

No change

Comment 3.
There are some serious methodological flaws in authors' approach to combining studies in conventional and network meta-analysis. First, NLST is essentially a head-to-head lung cancer screening trial and should not be combined with other LDCT screening trials with usual care or no screening as comparison.

Response 3.

NLST claims CXR screening is usual care. While this claim is debatable, NLST is the trial which is being used to support the introduction of lung cancer screening with LDCT in health services where no screening is undertaken. It thus seems perverse to exclude the most important piece of evidence being used to justify screening. Further this is easily achieved if the review question is LDCT versus any comparator, which is what our review question was. We used network meta-analysis to address the effect of variation in comparators. We therefore dispute the contention that this is a methodological flaw.

Change 3.

No change

Comment 4.

There are some serious methodological flaws in authors' approach to combining studies in conventional and network meta-analysis. …… Secondly, MILD study is essentially two trials in one i.e. Annual screening arm and Biennial screening arm and study reports very different results in terms of magnitude of effect on mortality for each when compared with usual care, so combining them in to a single comparison is methodologically not correct.

Response 4.

This is clearly understood and recognized in our description of the included studies.

Given that we are uncertain about the presence of an effect at all, it seems premature to assume that there is a dose effect, and what this might be.

It is important to note that the approach we used was aimed for the avoidance of the unit-of-analysis error which would occur in a situation where a three-arm trial was included. Therefore, our approach has avoided the unit-of-analysis error for a study that could contribute multiple, correlated, comparisons. This is an approach recommended by the Cochrane Handbook.

For completeness, we did repeat the analyses formally separating the two arms of MILD, but this made very little difference to the summary results, as would be anticipated from the extremely low weighting given to the MILD study 6.58/7%
Comment 5.

To properly answer the question on effect of LDCT screening on lung cancer mortality, one has to account for subsequent harms associated with screening as well such as over diagnosis, false positives and consequences of false positives, so an informed decision on effectiveness of screening can be made i.e. balancing benefits vs. potential harms.

Response 5.

We believe that it is still useful to look at outcomes in isolation, particularly where they are as important as mortality. This has allowed us to look at the data in more detail than possible when trying to examine all outcomes. We think this focus has been useful and uncovered complexities which have been glossed over previously. We respect that papers looking across all outcomes are useful too; we are just clear that this is not what we are doing in this paper.

Comment 6.

Apart from doing risk of bias assessment for individual studies, authors have refrained from doing an actual quality assessment of body of evidence i.e. GRADE. The pooled estimates have little to no certainty if they are being informed through poor quality of evidence.

Response 6.

Given that we are focusing on single outcomes, there is no need to GRADE the body of evidence. We do go in to quality assessment of the trials in some detail however, including a detailed analysis of the base-line characteristics.

Comment 7.
Authors also did not provide additional information on important subgroups of interest where a differential effect of screening may be present such as stage of lung cancer.

Response 7.

Looking at sub-groups would have been nice, but this did not seem justifiable given the limited power of the whole trials to address mortality and the limited information contained about subgroups in the trial reports. We did not pre-specify any sub-groups so the results of such an analysis would be post hoc and hypothesis generating only. Given this we felt the added value of doing sub-group analyses in response to the peer reviewer was minimal.

Change 7.

No change

Reviewer #2:

General comment. Thank you for the opportunity to review this interesting and concisely written manuscript. Although I have some concerns about the presentation of results (e.g. p-values, cross-tables) and description of methods (i.e. what model was used for network meta-analysis) I believe the review is comprehensive. The conclusion that, on balance, we should wait for further evidence seems justified if considering only the narrow grounds of preventable deaths, rather than wider effectiveness and cost-effectiveness.

Specific comments.

Comment 8.

The authors repeatedly refer to statistical significance but don't report any of the p-values. The recent American Statistical Association statement on p-values advises against the use of the arbitrary 5% threshold for decision making and that p-values themselves should always be reported (American Statistical Association 2016). Please include these in the abstract and wherever in the text a statement on significance is being made. Ideally, the authors would instead refer to weak or strong evidence in support of hypotheses(Sterne and Davey Smith 2001).

Response 8.

Thank you for your comment and we have added p-values in the abstract and the main text.

Change 8
Exact p-values added

Comment 9.
On page 5 the authors present the methods for network meta-analysis (NMA) but do not specify the link function or likelihood. The National Institute for Health and Care Excellence provide guidance on the appropriate model for binomial outcomes where follow-up varies across trials (i.e. binomial likelihood and complementary log log link). As the authors present risk ratios, and not hazard ratios, it is unlikely the NICE TSD model was adopted. Please be clear what model was used.

Response 9.
Thank you for your comment and we have added the following to the methods section:

Multivariate random-effects meta-analysis model using restricted maximum likelihood approach was performed.

Change 9.
Change as indicated

Comment 10.
On page 8 is reported the confidence interval for LDCT screening vs controls on all-cause mortality (0.89 to 1.00). This is interpreted as "not-statistically significant" and is dismissed. This illustrates the danger of using the arbitrary 5% threshold for decision making as this is only marginally non-significant and really provides weak evidence that LDCT screening reduces all-cause mortality.

Response 10. Agreed
Change 10. Have added comment to note borderline nature of the result’s statistical significance

Comment 11.
On page 8 specific risk ratios estimated by the NMA are presented. Please provide a cross-table with all pairwise comparisons of screening methods on both all-cause and lung cancer mortality.

Response 11.
Thank you for your comment and we have added tables for the risk ratios estimated by the NMA in the appendix.

Change 11.

See Web tables 5 & 6

Comment 12.

Page 9 of the discussion gives an interesting discussion of numbers needed to screen to avoid one lung cancer death (which is 357). This would be boosted by giving the proportion of patients in the high risk group and proportion likely to turn up for screening. A mention of the actual cost of screening would provide further useful context.

Response 12.

We considered this invitation carefully. We decided against including information on the proportion likely to turn up to screening because this is currently not known. The included RCTs do not provide this information as the trial participants are volunteers – although compliance in these was good. We decided against including information on cost, as this is too complex an issue to deal with in a simple statement. For instance LDCT price may depend on volume. Similarly whether the most useful cost would be of the screening test alone, or the nett cost of the screening programme, taking into account the cost of follow-up tests, was not obvious to us.

Change 12.

No change

Comment 13.

Cost-effectiveness modelling is only mentioned at the end of the conclusions. The findings of this analysis (now published, and to the effect that there is evidence a single round of screening could be cost-effective) should be reported and discussed. It would follow on naturally from the discussion on numbers needed to screen.

Response 13.

Again we considered this invitation carefully. On balance we felt that the message on cost-effectiveness was sufficiently complex that it was not sensible to summarise it divorced from the analysis which gave rise to the conclusions. We adopted a particular modelling approach and used assumptions which need to be appreciated in order to properly understand the cost-effectiveness estimates we produced. We were actually sceptical about the cost-effectiveness of
LDCT screening because optimistic estimates of effect on mortality tend to have been used in cost-effectiveness models.

We introduced the number needed to screen estimate mainly to emphasize that the relative effect exaggerates the absolute change in risk because the baseline risk of lung cancer is low, rather than as a route to a simplified summary of benefits, harms and costs.

Change 13.

No change

*** Comment 14.

We have added an additional author, Jaime Peters, to recognize the assistance that they provided in responding to comments.