

# **Public Health England External Review of the AgeX Trial**

**p1-4 Reviewer comments (mean score = 9.3/10)**  
**p5-7 Investigators' replies to queries by reviewers**

February 2017



Professor Dame Valerie Beral  
Director of the Epidemiology Cancer Unit  
Nuffield Department of Population Health  
Medical Sciences Division  
Oxford University

[pa.valerie.beral@ceu.ox.ac.uk](mailto:pa.valerie.beral@ceu.ox.ac.uk)

9 February 2017

Dear Valerie

### **External reviews of AgeX – a summary of findings**

As discussed, herewith the summary of findings from the external reviews of AgeX.

#### Introduction

PHE consulted three external academics to review the revised AgeX protocol with regards to its scientific merit. All were asked some general questions and then asked to provide some constructive critical feedback.

#### Conditions

All reviewers were informed that their identities and affiliations would remain anonymous but that their opinions would be shared with the investigators.

#### Results

1. All three reviewers gave positive responses to the following questions which queried the research question, governance and basic premise of the Trial.

Is the Trial:

- Of public health value
- Feasible study
- Satisfactory track record of investigators
- Ethics committee and governance approval
- Patient/public involvement

2. The reviewers were asked to score the Trial on three criteria. Table 1 shows the scores of the 3 individual reviewers for each criterion. A score of 9-10 was deemed excellent / essential; 7-8 was rated very good/highly relevant/highly realistic.

| <b>Scoring criteria</b>                          | <b>Reviewer<br/>1</b> | <b>Reviewer<br/>2</b> | <b>Reviewer<br/>3</b> |
|--|-----------------------|-----------------------|-----------------------|
| Scientific merit                                 | 9                     | 10                    | 8/9                   |
| Research methods, including recruitment          | 9                     | 10                    | 9                     |
| Research team, resources and research management | 10                    | 10                    | 8/9                   |

3. Reviewers were asked to comment on the key strengths of the proposal. Responses received included:
- Cluster randomisation
  - Strong scientific team
  - Appropriate outcomes
  - Adequate statistical power
  - The best scientific method to establish the effect of adding one breast screen before age 51, and of adding breast screens after age 69. Extremely large sample size, and therefore power.
  - This ongoing pragmatic trial of extending mammographic screening to women younger than 50 and older than 70 is crucial for providing evidence on the benefits and risks of screening in the extended age groups. The design of this trial is novel in its pragmatic design and fully embedded in the ongoing UK Breast Cancer Screening Program. The trial will provide evidence on the value of expansion of the screening age range in the UK universal health care system, within the context of contemporary breast cancer screening and treatment methods, which is a major strength. The primary outcome of the trial, breast cancer mortality, is an appropriate metric for assessing the value of mammographic screening. In the absence of any benefit (if that is what emerges), the trial is well positioned to evaluate disbenefits to women and costs to the healthcare system of unnecessary screening. The ability to exclude women, based on data prior to randomization, who would not be responsive to an invitation to breast cancer screening, is also a strength that improves power a bit. Because of these many strengths, the evidence generated from this study will not only benefit women's public health in the UK, but worldwide. We eagerly await the results of this important trial.
4. General comments and constructive feedback were invited by the reviewers. These are illustrated in sections a-h.
- a. Selective invitations

*I am pondering the design feature of excluding women who would be unlikely to respond to the screening invitation based on past behaviour. I understand that the key question is one of efficacy (i.e., evidence for benefit under optimal circumstances). It would seem desirable to analyse the outcomes from multiple perspectives – intent to treat, likelihood of participating (which will exclude some who participate and include some who won't), and actual participation to see where/if there is convergence.*

b. Assessing both harms and benefits

*A second comment is to prioritize assessment of harms as well as benefits. Evidence-based guidelines in other countries that might derive from these studies need to be informed by the net benefit of additional early or later screening.*

c. Contamination from additional screenings

*Are women in the trial able to obtain mammographic screening outside of the UK Breast Cancer Screening Programme, from private providers? Is this a possible source of cross-over that could affect, in any major way, the results particularly among women with additional access to private healthcare?*

d. Is it time for an interim analysis?

*Is it still ethical to randomise women for screens after age 69, or a control arm not receiving breast screens, now that IARC with several international experts has established that there is sufficient evidence for breast screening women aged 70-74?  
Or in other words, would it be a suggestion to do an interim analysis for this age group, perhaps earlier than expected? Note, the BC mortality reductions in the older age groups are suggested to be large.*

e. Query on mortality reduction estimate

*Although the sample size is exceptional, the assumption of a 15% BC mortality reduction of just one screen before age 51, seems perhaps slightly high (also considering no contamination). Could investigators address on which this has been based, and whether with more recent knowledge one would still find this to be the correct assumption?*

f. Follow-up and metastases

*The death review is done by narratives based on the 1<sup>st</sup> diagnosis. Is there no information in the narratives about follow up, and metastases (treatment), etc?*

g. Exclusion criteria

*Why is there exclusion of cases based on “breast disease / surgery” only?*

h. Length of follow-up

Could investigators explain a little bit more about the limitation of deaths up to age 60, or age 80? Is that not a too small age range/follow up period?

Conclusion

The three external reviewers all support the trial methodology and consider it has much scientific merit. PHE would welcome the views of AgeX investigators on some of the questions raised in the general comments section. This could help inform further debate on the AgeX trial.

I hope this is helpful.

Yours sincerely



---

Anne Mackie  
Director of Screening – PHE  
[anne.mackie@phe.gov.uk](mailto:anne.mackie@phe.gov.uk)

cc Professor Kevin Fenton  
Director, Health and Wellbeing  
PHE

# Reply by Investigators to Public Health England about Queries Raised by External Reviewers

The queries (a-h) raised by the Public Health England (PHE) external reviewers are shown in italics (in red) and are followed by the investigators' replies

*Query a. Selective invitations: I am pondering the design feature of excluding women who would be unlikely to respond to the screening invitation based on past behaviour. I understand that the key question is one of efficacy (ie, evidence for benefit in optimal circumstances). It would seem desirable to analyse the outcomes from multiple perspectives – intent to treat, likelihood of participating (which will exclude some who participate and include some who won't), and actual participation to see where/if there is convergence.*

**Reply:** Trial protocols should specify unambiguously what the principal analysis will be, and to avoid bias it is usual for this to involve analysis by allocated treatment (which also known as intent-to-treat analysis – the two terms are exactly equivalent). So, the principal analysis cannot be of the women who actually participate vs the women who do not (as this could well be seriously biased by the “healthy volunteer” effect).

To maximize statistical sensitivity, the principal intent-to-treat analysis in AgeX excludes women who, based unbiasedly on information recorded before randomisation, would be unlikely to accept a screening invitation (eg, women who ignored their previous breast screening invitation would be unlikely to accept an invitation to be screened in AgeX, so such women are excluded from the principal intent-to-treat analysis of women who, if invited, would be likely to accept.) This improves the statistical power of the study.

## **Quotations:**

“The ability to exclude women, based on data prior to randomization, who would not be responsive to an invitation to breast cancer screening, is also a strength that improves power...” [page 2 of the report of the PHE reviewers]

“The principal analyses will be restricted to those women among whom, based on information recorded prior to the random allocation, an invitation would be likely to have made them attend for screening if they would not otherwise have done so. Among them, analyses by allocated treatment will be used...” [page 1 of AgeX trial protocol; see also section on page 6, Exclusions from primary analyses and main subsidiary analyses]

*Query b. Assessing both harms and benefits: A second comment is to prioritize assessment of harms as well as benefits. Evidence-based guidelines in other countries that might derive from these studies need to be informed by the net benefit of extra early or later screening.*

**Reply:** Ever since the study was planned this has been its major aim, so although the principal analysis is of any effects on breast cancer mortality, the main subsidiary analyses are of any harms arising from over-diagnosis and over-treatment.

## **Relevant quotations:**

“This trial will assess the risks of screening (in particular, the chances of being diagnosed and treated for a non-life-threatening cancer) and benefits (in particular, the chances of saving life)” [AgeX patient information sheet]

“The main subsidiary analyses will be of the details of breast cancer incidence and of the patterns of breast cancer investigation and treatment. Information on screening outcomes, such

as recall and biopsy rates, will be collected not only for the women randomised to extra screening invitations but also for the first routine screening invitations at ages 50-52. In addition, many other outcomes available from linkage with routine NHS records will be assessed.” [page 6 of AgeX trial protocol]

*Query c. Contamination from additional screenings: Are women in the trial able to obtain mammographic screening outside of the UK Breast Cancer Screening Programme, from private providers? Is this a possible source of cross-over that could affect, in any major way, the results particularly among women with additional access to private healthcare?*

**Reply:** They can do so, and some will do so, but the numbers doing so at present are too small to affect the findings in any major way.

*Query d. Is it time for an interim analysis? Is it still ethical to randomise women for screens after age 69, or a control arm not receiving breast screens, now that IARC with several international experts has established that there is sufficient evidence for breast screening women aged 70-74? Or, in other words, would it be a suggestion to do an interim analysis for this age group, perhaps earlier than expected? Note, the breast cancer mortality reductions in the older age groups are suggested to be large.*

**Reply:**

Interim analyses of both harms and benefits will be provided at least annually for the Data Monitoring and Ethics Committee (and more frequently if DMEC considers it necessary), and their terms of reference are to advise if there is ever proof beyond reasonable doubt that screening invitations would be appropriate for some or all women. **At its 20 Feb 2017 meeting to review confidential mortality and recurrence analyses, the DMEC did not consider there to be any need to request more analyses or disclose interim findings.**

**Quotation:**

“The data monitoring and ethics committee, which is independent of the trial team, will oversee safety, efficacy and ethical issues, including any that arise from new information from other sources. It will confer no less than about once a year, and can request extra meetings at any times it considers appropriate. Progress reports and data will be provided when it confers, and it can demand any analyses or information it considers appropriate to inform its decisions.” [page 9, AgeX trial protocol]

*Query e. Query on mortality reduction estimate: Although the sample size is exceptional, the assumption of a 15% breast cancer mortality reduction from just one screen before age 51 seems perhaps slightly high (also considering no contamination). Could investigators address on what this has been based, and whether with more recent knowledge one would still find this to be the correct assumption?*

**Reply:**

**Recent knowledge lets the primary analysis detect a mortality reduction of only 10% with the sensitivity that with earlier knowledge was possible for a 15% reduction.**

By early 2017, we have cancer registration data for women who entered the study at ages 47-49 and who were followed for cancer for at least 4 years. Half these women were allocated an extra screening invitation at study entry and half were not, but both groups were invited for screening 3 years later. By 4 years after study entry 1.1% of both these groups had had a breast cancer diagnosed, as the screening invitation for all the women 3 years after entering the study resulted in complete catch-up of the cumulative number of cases in those who had not been given an extra invitation when they originally entered the study.

Because catch-up is complete by year 4, there will be no material differences between the two groups after year 4 in the rates of diagnosis of breast cancer, and breast cancers arising more than 4 years after study entry are uninformative about the effects of the extra screening on the breast cancers diagnosed in the first 4 years after study entry. This means that for women who enter the study at ages 47-49 the primary analysis that is specified by the protocol will, without introducing any material bias, become an analysis of death from a breast cancer that was diagnosed within 4 years of study entry. This will reduce by more than half the random variation in the main findings. Even without eliminating the uninformative breast cancers the study could reliably detect a 15% reduction in breast cancer mortality, but now it can just as reliably detect a 10% reduction.

**Quotation:**

“In younger women deaths from breast cancer diagnosed after the first routine screen at age 50-52 would not be expected to be affected by the random allocation and hence will be uninformative. To achieve greater sensitivity, the primary analyses will consider separately these uninformative breast cancer deaths and all other breast cancer deaths, if this can be done reliably without introducing any material bias...” [page 6, AgeX trial protocol]

*Query f. Follow-up and metastases: The death review is done by narratives based on the first diagnosis. Is there no information about follow up, metastases, treatment, etc?*

**Reply:**

See the middle two paragraphs on page 5 of the protocol, which state that to determine whether uncontrolled life-threatening breast cancer was present at the time of death the Endpoint Committee will review the deaths of all women with any history of breast cancer. Their review will be based on access to all relevant information from hospital admissions throughout the course of the disease that is provided electronically by HES. This would include information about the treatments given and the development of metastases.

**Quotation:**

“For each probable breast cancer death, trial organisers will seek a narrative of the diagnosis and treatment of that cancer.” [page 5, AgeX trial protocol]

*Query g. Exclusion criteria: Why is exclusion of cases based only on “breast disease/surgery”?*

**Reply:**

Exclusion is based on many things, of which this is only one: see the 3-paragraph section spanning pages 6-7 of the AgeX trial protocol, entitled “Exclusions”.

*Query h. Length of follow-up: Could investigators explain a little bit more about the limitation of deaths up to age 60, or age 80? Is that not a too small age range/follow up period?*

**Reply:**

This would indeed be too small an age range/follow-up period, but where these limitations are stated, the protocol also states that eventually follow-up will continue at least 15 years.