Sussex Research Online

Ovarian cancer population screening and mortality after long-term follow-up in the UK collaborative trial of ovarian cancer screening (UKCTOCS): a randomised controlled trial

Article (Accepted Version)

Menon, Usha, Gentry-Maharaj, Aleksandra, Burnell, Matthew, Singh, Naveena, Ryan, Andy, Karpinskyj, Chloe, Carlino, Giulia, Taylor, Julie, Massingham, Susan K, Raikou, Maria, Kalsi, Jatinderpal K, Woolas, Robert, Manchanda, Ranjit, Arora, Rupali and Fallowfield, Lesley (2021) Ovarian cancer population screening and mortality after long-term follow-up in the UK collaborative trial of ovarian cancer screening (UKCTOCS): a randomised controlled trial. Obstetrical and Gynecological Survey, 76 (9). pp. 537-538. ISSN 0029-7828

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/106431/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

General population screening and ovarian cancer mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial

Prof Usha Menon FRCOG [1], Aleksandra Gentry-Maharaj PhD[1], Matthew Burnell PhD[1], Prof Naveena Singh FRCPath[2], Andy Ryan PhD[1], Chloe Karpinskyj MSc[1], Giulia Carlino MSc[1], Julie Taylor MSc[3], Susan K Massingham RGN[1], Maria Raikou PhD[4], Jatinderpal K Kalsi PhD[5], Robert Woolas FRCOG[6], Prof Ranjit Manchanda PhD[7], Rupali Arora FRCPath[8], Laura Casey FRCPath[2], Anne Dawnay PhD[9], Stephen Dobbs FRCOG[10], Simon Leeson FRCOG[11], Tim Mould FRCOG[12], Mourad W Seif FRCOG[13], Aarti Sharma FRCOG[14], Karin Williamson FRCOG[15], Yiling Liu MS[16], Prof Lesley Fallowfield DPhil[17], Prof Alistair J McGuire PhD[4], Prof Stuart Campbell FRCOG[18], Steven J Skates PhD[16], Prof Ian J Jacobs FRCOG[5,19], Prof Mahesh Parmar DPhil[1].

[1]MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College London WC1E 6AU, London, UK
[2] Department of Cellular Pathology, Barts Health NHS Trust, London E1 1BB, London, UK
[3]Clinical Epidemiology, Institute of Health Informatics, University College London, London NW1 2DA, London, UK
[4] London School of Economics and Political Science, London WC2A 2AE, London, UK
[5] Department of Women's Cancer, Institute for Women's Health, University College London, London WC1E 6HU, London, UK
[6] Department of Gynaecological Oncology, Queen Alexandra Hospital, Portsmouth PO6 3LY, Portsmouth, UK
[7] Barts Health NHS Trust and Wolfson Institute of Preventive Medicine, CRUK Barts Cancer Centre, Queen Mary University of London, London E1 1BB, London, UK
[8] Department of Cellular Pathology, University College London Hospitals NHS Trust, London W1T4EU, London, UK
[9] Department of Clinical Biochemistry, Barts Health National Health Service Trust, London E1 2ES, London, UK

[10] Department of Gynaecological Oncology, Belfast City Hospital, Belfast BT9 7AB, Belfast, UK
[11] Department of Obstetrics and Gynaecology, Ysbyty Gwynedd, Bangor, LL57 2PW, UK
[12] Department of Gynaecological Oncology, University College Hospital, London NW1 2BU, London, UK
[13] Divisions of Gynaecology and Division of Cancer Sciences, St. Mary's Hospital, Manchester M13 9WL, Manchester, UK
[14] Department of Obstetrics and Gynaecology, University Hospital of Wales, Cardiff, CF14 4XW, UK
[15] Department of Cancer Services, Royal Free Hospital, NW3 2QG, London, UK
[16] Massachusetts General Hospital and Harvard Medical School, Boston MA 02115, Boston, USA
[17] Sussex Health Outcomes Research and Education in Cancer (SHORE-C), Brighton and Sussex Medical School, University of Sussex, Brighton BN1 9RX, Brighton, Sussex
[18] Create Health, London EC2V 6ET, London, UK
[19] Department of Women's Health, University of New South Wales, NSW 2052, Sydney, Australia

Corresponding Author
Professor Usha Menon
MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology
University College London
90 High Holborn, 2nd Floor, London WC1V 6LJ
+44 (0)20 7670 4649 u.menon@ucl.ac.uk

Brief title: Long term impact of screening on ovarian cancer mortality

Key words: Ovarian cancer, screening, UKCTOCS, mortality, randomised controlled trial, RCT, CA125, TVS, long term

Summary

Background We report on ovarian cancer mortality after long term follow-up in the UKCTOCS.

Methods UKCTOCS is a 13-centre randomised controlled trial of postmenopausal women from the general population, aged 50–74, with intact ovaries, no ovarian or active non-ovarian cancer or familial ovarian cancer risk. Participants were randomly allocated 1:1:2 by the trial management system to annual screening till Dec 31, 2011 using a multimodal (MMS, 345570 screens) or ultrasound (USS, 327775 screens) strategy or no screening. Follow-up was through national registries. Primary outcome was death due to ovarian/tubal cancer (WHO2014) by June30, 2020. Analyses were by intention to screen, comparing MMS and USS separately with no screening using the Versatile test. Investigators and participants were aware and outcomes review committee were masked to randomisation group. ClinicalTrials.gov registrationNCT00058032.

Findings Eligible women were 202562: 50625MMS, 50623USS and 101314no screening group. At a median follow-up of 16·3(IQR15·1-17·3) years, 2055 women developed tubal/ovarian cancer: 522 (25·4%) MMS, 517 (25·2%) USS, 1016 (49·4%) no screening. Compared to no screening, there was a 47% (95%CI:19·7,81·1) increase in stage I and 24·5% (95%CI:-41·8,-2) decrease in stage IV incidence in the MMS group. Overall the incidence of Stage I/II was 39·2% (95%CI:16·1,66·9) higher and Stage III/IV 10·2% (95%:-21·3,2·4) lower in the MMS group. 1206 women died of the disease: 296(24·6%) MMS, 291(24·1%) USS, 619(51·3%) no screening. There was no significant reduction in ovarian/tubal cancer deaths in either MMS(p=0·580) or USS(p=0·360) compared to no screening group.

Interpretation The reduction in stage III/IV incidence in the multimodal group was not sufficient to translate into lives saved illustrating the importance of specifying cancer mortality as the primary outcome in screening trials. Given that screening did not reduce ovarian/tubal cancer deaths, currently general population screening cannot be recommended.

Funding

National Institute for Health Research, Cancer Research UK, The Eve Appeal.

INTRODUCTION

Ovarian cancer remains the most fatal of all gynaecological cancers. The majority (58%) of women are diagnosed at an advanced stage (III/IV) which is associated with poor survival (5-year survival 27% Stage III, 13% Stage IV).[1] The over 90% survival rates in women detected at StageI[1] has spurred international efforts in early detection, spanning over four decades.[2-6] All trials have used combinations of the biomarker CA125 and/or pelvic imaging using transvaginal ultrasound scan (TVS). Despite these extensive endeavours, to date there is no evidence that screening for ovarian cancer saves lives.[7-9]

In our multicentre randomised trial (UKCTOCS), 202638 women from the general population were randomised to two screening groups – multimodal (MMS) and TVS (USS) and a 'no screening' group. We reported earlier (median follow-up of 11·1 years), that an absolute of 13% more women with ovarian, tubal and peritoneal cancer were diagnosed with Stage I/II disease in the MMS ($p<0·001$) compared to 'no screening'. There was no change in stage in the USS group [9] Despite this, there was no evidence of a reduction in disease specific deaths in either screened group compared to the 'no screening' group using the Cox version of the logrank test. The observed reduction in deaths was delayed and the cumulative mortality curves appeared to be diverging at time of previous reporting.[9] We therefore continued follow-up and report here on the long-term mortality impact of ovarian cancer screening in UKCTOCS.

METHODS

Study design and participants

UKCTOCS is a randomised controlled trial (RCT) of 202638 postmenopausal women aged 50-74 from the general population. UKCTOCS was approved by the UK NorthWest MREC (00/8/34) on June23,2000. The trial design has been previously published and the protocol is available online.[9-12]

In brief, we invited 1243282 women from age sex registers of Primary Care Trusts in England, Wales and Northern Ireland and randomised 202638(16·3% of 1243282)

through 13 trial centres based at adjoining National Health Service (NHS) Trusts between April17,2001 and Sept29,2005.[10] Inclusion criteria were 50–74 years of age and postmenopausal status. Exclusion criteria were bilateral oophorectomy, previous ovarian or active non-ovarian malignancy or increased familial ovarian cancer risk.

Randomisation and masking

The trial management system confirmed eligibility and then randomly allocated women using the Visual Basic randomisation statement and the Rnd function: 50640(25·0%) to MMS, 50639(25·0%) to USS, and 101359(50·0%) to no screening. Investigators and participants were aware and the outcomes committee was masked to randomisation group.

Screening

Annual screening in the MMS group used serum CA125 measurements, with the pattern over time interpreted using the risk of ovarian cancer (ROCA) calculation,[13] which identifies significant rises in CA125 concentration above baseline. Based on risk women were triaged to normal (annual screening), intermediate (repeat CA125 ROCA test in 3 months), and elevated (repeat CA125 ROCA test and transvaginal USS as a second-line test in 6 weeks) risk. Annual screening in the USS group used TVS as the primary test, which was classified as normal (annual screening), unsatisfactory (repeat in 3 months), or abnormal (scan with a senior ultrasonographer within 6 weeks). Screening ended on Dec31, 2011. We undertook 673345 annual screens; 345570 MMS group, 327775 USS group. Compliance with screening was high (MMS81%; USS78%) with women undergoing a median of 8 annual screens.[9] In both groups, women with persistent abnormalities were assessed by a trial clinician and underwent further investigation within the NHS. We deemed women who had surgery or a biopsy for suspected ovarian cancer after clinical assessment as screen positive.

Follow-up

Women were linked using their NHS number to national cancer and death registration data and hospital episodes administrative records. Additional sources included cancer diagnosis data from the National Disease Registration Service for English participants and postal follow-up questionnaires; two prior to previous

analysis (3-5 years after randomisation, and April2014) and a third during long-term follow-up (appendix p 4). The third questionnaire was sent in June 2020 to a subset of 6586 participants who had either exited the national registries or for whom it was not possible from HES data to ascertain if both ovaries had been removed. We also had ad-hoc direct communication from trial participants, their families and physicians.

Cancer site and cause of death review

Throughout the trial, we interrogated the above sources to identify women diagnosed with any of 19 International Classification of Diseases(ICD)-10 codes as detailed in UKCTOCS protocol and retrieved copies of medical notes.[12,14] The only exception was women with malignant neoplasm of uncertain origin (ICD-10 C80) who also had another non-ovarian, tubal or peritoneal cancer registration. Notes with any reference to randomisation group redacted were reviewed by the outcomes review (OR) committee consisting of gynaecological pathologists and oncologists. The OR committee assigned cancer site (using a previously audited pre-specified algorithm),[14] FIGO 2014 stage at diagnosis, grade, morphology, type of ovarian cancer and cause of death. We defined ovarian/tubal cancer using the revised WHO2014 classification[15,16] and death due to ovarian and tubal cancer based on disease progression (appearance of new lesions or increases in size of previously documented lesions with imaging, clinical worsening, or rising biomarker concentrations). In the WHO2014 classification, the definition for primary peritoneal cancers was revised. The OR committee Chair therefore reviewed all 41 cancers previously classified as primary peritoneal as per WHO2003 classification.[9] The OR committee re-staged using FIGO2014 criteria all ovarian/tubal cancers diagnosed in 2001-2014 as they had been previously staged using FIGO 2003 criteria.

Outcomes

The primary outcome was death due to ovarian (ICD-10 C56) or tubal (ICD-10 C57·0) cancer. Ovarian cancer includes primary non-epithelial ovarian cancer, borderline epithelial ovarian cancer, and invasive epithelial ovarian cancer. As stated above ovarian cancer was defined using the revised WHO2014 definition.[15,16] This is in contrast to the previous mortality analysis (censorship Dec 31, 2014) which used WHO2003 criteria.[17] The majority (40/41) of previously classified primary peritoneal

cancers using WHO 2003 criteria were reclassified as ovarian or tubal cancers. Secondary outcomes were ovarian/tubal cancer (1) incidence and (2) stage. For all outcomes, subgroup analysis was undertaken for invasive epithelial ovarian/tubal cancer. All outcome data were kept confidential until unblinding.

Power calculation

At previous analysis (censorship Dec31, 2014), there were 358 ovarian/tubal cancer deaths in the no screening group. Compared to the no screening group, the 'average' estimated relative mortality reduction (MR) in deaths was 11% (Cox model p=0·240) MMS and 9% (Cox model p=0·32) USS. Any MR was only apparent about 7 years after randomisation. 45% (162/358) of the deaths in the no screening group during 2001-2014 had occurred before 7 years. In 2015, for the no screening versus MMS or USS comparisons, we estimated that an additional 233 no screening group events would give 80% power at a two-sided 5% significance level for a difference in relative mortality of 25% during long-term (2015-2020) follow-up conditional on the observed MR of 11%. This translated to a target sample size of 591 overall events in the no screening arm. All 233 new and 73% (431/591) of total no screening group events would occur beyond 7 years. No formal adjustment was made to the test for a) having previously analysed the data in 2015 or b) making two screen group comparisons. Instead, we decided to openly describe the multiplicity issues and acknowledge the unadjusted p-values. As the number of events were less than anticipated on the planned censorship date of December31, 2018, it was extended to June30,2020.

Statistical analysis

Descriptive statistics regarding ovarian cancer death and incidence were created including tabulations of histology, stage and screen type by group.

The primary analysis was changed from that we used previously. In the 2015 report, we used a Cox version of the logrank test[9] which is most powerful under proportional hazards, to analyse the mortality data. For the current analysis, we extensively discussed the best approach within Trial Management and Trial Steering Committees, and consulted 12 independent international statistical, trial and screening experts. The details and rationale underpinning this important change are

reported separately.[18] In short, given the accumulating external evidence of delayed mortality effects in screening trials, the majority of the experts supported the change in primary analysis to a test that was sensitive to delayed effects. We felt it was important to choose a test that was agnostic to the specific form of the screening effect. We therefore chose the Versatile test described in 2016[18] that is a combination test of 3 log-rank test statistics (Z_1, Z_2, Z_3), covering early, constant, and late effects respectively (further details appendix p3).

All analyses were by intention to screen. The primary mortality analysis was an MMS versus no screening and USS versus no screening analysis of the primary outcome using the Versatile test,[19] with a Royston-Parmar (RP) model[20] used to estimate survival differences. We defined survival time from date of randomisation (t_0=0) to date of death due to ovarian/tubal cancer or censorship, or sooner if the volunteer died of another cause or was lost to follow-up. No allowances were made for non-compliance to screening or contamination (ovarian cancer screening) in the no screening group. We describe potential time-dependent features of the screening effect by estimating the hazard ratio and the absolute survival difference at the pre-specified time-points of t=5, 10, 15 and 18 years (maximal follow-up was 19·3 years) using a flexible parametric RP model[20] (appendix p3).

We undertook two secondary analyses of the primary outcome. We fitted a proportional hazards Cox model to the primary outcome data. To allow for a formal analysis of the late effects of ovarian cancer, not subject to issues of data re-use[21] and multiple testing, we also fitted a Cox model to the 'new' data acquired since 1 January 2015. Both the methods and results of sensitivity analyses are detailed in the supplementary materials (appendix pp8,9). Survival from diagnosis in women with ovarian and tubal cancer in the no screening group was also compared to national age and period adjusted one, five, and 10-year survival rates. We undertook a subgroup analysis using the Versatile test of invasive epithelial ovarian and tubal cancer death where other ovarian cancers were censored at death.

For the secondary outcome, cumulative incidence of ovarian/tubal cancer were presented graphically using standard Kaplan Meier (KM) methods, based on time from randomisation to diagnosis. Death from other causes and bilateral

oophorectomy were censoring events. Administrative censorship was the same as for the mortality analysis (June30, 2020). Ovarian and tubal cancer incidence rates were explored parametrically using a RP model. For the secondary outcome of ovarian and tubal cancer incidence by stage, and the subgroup analysis of invasive epithelial ovarian and tubal cancers, we used incidence rate ratios (IRR) with 95% CI[22] to compare no screening versus MMS and USS groups separately. We also calculated stage-specific ovarian/tubal cancer case fatality rates.

We used Stata(version16) and R(version 4.0.2) for all statistical analyses This trial is registered with ISRCTN number 22488978; ClinicalTrials.gov. number NCT00058032.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. UM and AR extracted the dataset. MB, SJS, YL, AM, MR, UM, and AR had full access to the dataset. UM had final responsibility for the decision to submit for publication.

RESULTS

The final cohort eligible for analysis consisted of 202562(>99·9%) of 202638 women: 50625(>99·9%) in the MMS group, 50623(>99·9%) in the USS group, and 101314(>99·9%) in the no screening group. We excluded 76(<0·5%) women (MMS 15 [<0·5%]; USS 16 [<0·5%]; no screening 45 [<0·5%]) (figure 1). As previously reported, baseline characteristics were balanced between the groups.[9]

Following the end of annual screening on December31, 2011, all women were followed-up until the censorship date June30, 2020. Complete follow-up until censorship or death was possible in 192478 (95·0%) women (MMS 48 110 [95·0%]; USS 48022 [94·9%]; no screening 96276[95·0%]) resulting in 3·16 million women years. Median follow-up was 16·3 (IQR 15·1–17·3) years for all groups.

We identified 4482 women with the 19 pre-specified ICD-10 codes who were reviewed by the OR committee (appendix p 5). Of them, 2055 (46%) women were confirmed to have ovarian/tubal cancer (table 1). The incidence of ovarian and tubal

cancer (per 100 000 women-years) was 67·7(95%CI:61·9,73·5; 522 cancers; 770967 women-years) in the MMS group, 68·2(95%CI:62·4,74·1; 517 cancers; 755677 women-years) in the USS group, and 65·4(95%CI:61·4,69·4; 1016 cancers; 1552703 women-years) in the no screening group. Figures 3a and 3b provide KM cumulative cancer rates for all and for invasive epithelial ovarian and tubal cancers, respectively. Both plots show a greater number of cancer diagnoses in the screening groups during the first screening year reflecting the lead time to diagnosis achieved by screening. The difference was largely maintained throughout the screening phase before apparent catch-up by the no screening group during the extended period of follow-up after the end of screening. However, the pattern of catch-up in the USS group was less pronounced, and this is elucidated by the RP model hazard functions (appendix pp 19, 20) where the rate of cancer incidence drops below the no screening group between years 4-14 approximately, before rising back above the no screening group rate.

Overall 1805(452 MMS; 445 USS; 905 no screening) of 2055 women had invasive epithelial ovarian/tubal cancers. The proportion of Type II cancers (79·2% MMS; 82·2% USS; 76·4% no screening) was similar across the groups (table 2). At 9·5 years after end of screening, compared to the no screening group, there was a 47% (95%CI 19·7;81·1) higher incidence of stage I disease and a 25% (95%CI -41·8;-2) lower incidence of Stage IV disease in the MMS group (table 2). Overall, there was a 39% (95%CI:16·1,66·9) higher incidence of stage I/II and 10%(95%:-21·3,2·4) lower incidence of stage III/IV disease in MMS compared to no screening. For the subgroup analysis of invasive epithelial ovarian/ tubal cancers, the changes in stage distribution in the MMS group compared to no screening group persisted. There was no evidence of a change in incidence in any stage in the USS group compared to the no screening group.

At censorship, 1206(0·60%) women had died of ovarian cancer: 296(24·6%) MMS, 291(24·1%) USS, 619(51·3%) no screening group. Ovarian/tubal cancer deaths and incidence by year from randomisation is detailed in the appendix (appendix table 4, p7). The Versatile test (primary analysis) showed that there was no evidence of a reduction in ovarian/tubal cancer deaths in either the MMS(p=0·580) or USS(p=0·360) group compared to the no screening group (table 3). Figure 2a shows

the KM cumulative death rates; with any divergence between the screen and no screen groups being minimal. A sensitivity analysis that only considered data obtained by equivalent means of electronic health records across all three groups also showed no evidence of a difference using the Versatile test result for both the MMS(p=0·596) and USS(p=0·374) group. This and other sensitivity analyses are detailed in the appendix (appendix pp 8,9). A Cox model (secondary analysis) estimated a hazard ratio of 0·96(95%CI:0·83,1·10) for MMS versus no screening and 0·94(95%CI:0·82,1·08) for USS versus no screening. A Cox model fitted only to data from 2015 onwards (secondary analysis) estimated a HR of 1·05(95%CI:0·86, 1·30) for MMS versus no screening and 0·99(95%CI:0·80, 1·22) for USS versus no screening.

There were 295 deaths (0·58%) due to invasive epithelial ovarian/tubal cancer(secondary outcome) in the MMS group, 287(0·57%) in the USS group and 619 (0·61%) deaths in the no screening group. The cumulative death rates similarly showed no evidence of an effect of screening (Figure 2b). The Versatile test for mortality reduction showed no evidence of difference in both the MMS group (p=0·598) and the USS group (p=0·329). The RP model fit to the non-parametric KM curves (appendix pp 12,13) and the associated hazard functions for each group (appendix pp 14,15) are shown in the appendix. All hazard functions were monotonic, increasing with only small differences between the screen and no screen groups. At year 18 following randomisation, the RP model estimates of survival differences were 36·7(95%CI:-65·3,138·8) and 52·9(95%CI:-48·2,153·9) per 100000 women for MMS and USS compared to no screening, respectively (appendix p 10).

Compared to no screening, in the MMS group we observed a higher ovarian/tubal cancer case fatality rate in patients with Stage I disease and a lower mortality rate in Stage IV that persisted on subgroup analysis of invasive epithelial ovarian cancer. In the USS group, the stage specific rates were similar to the no screening group (table 2).

Survival from diagnosis in women with ovarian/tubal cancer in the no screening group was better in comparison to national age and period adjusted survival rates (1 year 77% vs 68%; 5 year 40% vs 37%) (appendix p 18).

DISCUSSION

Main findings

Our results from the largest ovarian cancer screening trial to date demonstrate that on long-term follow-up (median 16 years after randomisation), neither multimodal nor ultrasound screening, as used in UKCTOCS, significantly reduced deaths from ovarian and tubal cancer. There was a 47% higher incidence of Stage I and 24·5% lower of Stage IV resulting in an overall 39% higher incidence of Stage I/II and 10% lower of Stage III/IV in the multimodal group compared to the no screening group. General population screening for ovarian and tubal cancer with either of the screening strategies cannot be recommended. The changes in stage distribution during multimodal screening did not translate into mortality reduction, emphasising the importance of having disease-specific mortality as the primary outcome in cancer screening trials.

Achieving a mortality reduction will require a screening strategy that can detect ovarian/tubal cancer even earlier and in a larger proportion of women than we were able to achieve. Our findings make it even more critical that before general population screening is introduced, any new test is shown to reduce ovarian and tubal cancer deaths in a future RCT. These trials take many years to complete but the high compliance with annual screening in UKCTOCS suggests that women are very motivated to join them. Given that such trials take considerable time, it is likely that population screening for ovarian cancer is over a decade away.

It is difficult to extrapolate these results to ovarian cancer screening of high-risk women where the strategy has involved 3-4 monthly multimodal screening alongside risk-reducing surgery and resulted in a significant reduction in the proportion of women diagnosed with advanced disease.[23,24] There are also biological differences between cancers in women with BRCA gene mutations and the general population which result in improved treatment responses in the former. Unfortunately, it is unlikely that the true impact of screening on mortality will ever be assessed in this population as an RCT is challenging and there are potentially very effective

preventive measures such as risk reducing salpingectomy with delayed oophorectomy currently being evaluated.[25]

Our results have implications for ovarian cancer symptom awareness campaigns as the majority of women who were screen-detected had no 'high-alert' symptoms[26] and were diagnosed earlier than would have been possible with a symptom-based approach. This suggests that earlier diagnosis of invasive epithelial ovarian and tubal cancer in the symptomatic population is unlikely to translate into reduced mortality. However, it is important to note that there have been significant advances in treatment in the last ten years since the end of screening. The latter, in combination with earlier diagnosis may contribute to better quality of life and improved outcomes. In addition, achieving a rapid diagnosis is of great importance to women and their families.

In context of other trials

Our mortality results are similar to that reported in the ovarian arm of the PLCO screening trial, the only other large RCT to report on mortality impact. In the PLCO trial, there was no evidence of a reduction in ovarian cancer deaths between the screen and no screen arms, either at median follow-up of 12·4[7] or 14·7[8] years. However in UKCTOCS we found a higher incidence of Stage I and lower incidence of stage IV disease, in the multimodal group compared to the no screening group. The PLCO trial found no evidence of a difference in stage distribution between the screened and non-screened groups. The use in the MMS arm, of a longitudinal CA125 algorithm instead of a single CA125 cut-off as in PLCO trial may have contributed to this difference. We have previously shown that a longitudinal algorithm allows us to detect disease earlier and with greater sensitivity than a single CA125 cut-off.[27,28]

Despite the 24·5% reduction in stage IV disease in our multimodal group, the overall reduction in stage III/IV incidence 9·5 years after end of screening was only 10% with little change in Stage III incidence. Previous reports have highlighted the need for a large reduction in late-stage incidence as a prerequisite for reducing cancer mortality. However, it must be noted that the length of follow-up and therefore the dilution effect in the screen arms of inclusion of cancers diagnosed clinically after end of screening

varied in these reports. An analysis of breast cancer screening trials found no reduction in breast cancer mortality in trials that achieved less than 10% reduction in stage III/IV disease and an average reduction of 28% in trials that achieved a 20% or greater reduction.[29] In colorectal-cancer and lung cancer screening much of the mortality reduction is related to reductions in stage IV disease, which has much higher mortality than lower stages.[30,31] For ovarian/tubal cancer, the high mortality associated with stage III/IV, combined with most women clinically detected with stage III disease, requires a substantial reduction in the incidence of both stages before a mortality reduction is likely.

There are previous instances where increased incidence of stage I/II cancers in the screen group in screening trials did not translate to a mortality reduction. In the four early RCTs of lung cancer,[32-35] compared to the control group, the screened groups achieved significant improvements in stage distribution. However, much of the screen detected early stage disease in these trials were indolent. It was accompanied by a significant increase in cancer incidence in the screen groups, but no reduction in late-stage incidence. Both together suggest that overdiagnosis was the main contributor to the lack of reduction in mortality.[36] This differs from UKCTOCS where no significant increase in cancer incidence was observed in either of the screen arms. In the MMS arm of UKCTOCS it seems likely that the cancers shifted to an earlier stage had an intrinsic poor prognosis, which was not altered by earlier detection and the use of current stage based treatments. Further histopathological and genetic analysis could yield important information about the biology of ovarian cancer.

The 47% increase in incidence of stage I and 24% decrease in incidence in Stage IV was accompanied by a higher stage I and lower Stage IV case fatality rate in the multimodal group. This finding persisted on subgroup analysis of invasive epithelial cancers. The findings are unlike that described previously in cancer screening trials.[30] It suggests that in the MMS group, while earlier detection of Stage IV invasive epithelial ovarian/tubal cancers improved outcomes, earlier detection in Stage I of cancers that may have presented in later stages in the absence of screening, did not have the same impact. Stage specific mortality and treatment will

be the subject of further in-depth analyses that will be reported in a separate publication.

Strengths and limitations

Key strengths of UKCTOCS have been previously detailed[9] and include scale; multicentre design; adherence to protocol through use of a bespoke web-based trial management system with automation of key processes, remote data entry and concurrent central monitoring; completeness of follow-up through linkage to national registries and administrative databases and independent assignment of site and cause of death. The longitudinal algorithm we used to interpret CA125 levels was innovative and forward thinking. The UK government's new Accelerating Detection of Disease Programmme includes collection of repeat biological samples.[37] We re-staged all cases using the latest FIGO2014 criteria and revised our ovarian and tubal cancer site assignment using WHO2014 classification to reflect current understanding of disease biology. We also changed our primary analysis approach from a constant-effect approach (proportional hazards Cox-model) to one that allows for a delayed effect (the Versatile test) to reflect growing evidence that the mortality reduction in cancer-screening trials if present, is delayed. We did this through a transparent process with publication of our methods and the expert opinions that underpinned our decision.[18]

Much has been learnt from the design, conduct and analysis of UKCTOCS, that is relevant to future large-scale trials. In addition, a large bioresource of serum samples (>550 000) and linked data (UKCTOCS Longitudinal Women's Cohort, UKLWC) has been built through the generosity of the participants. It includes a unique sample set of up to 11 annual blood samples from women in the multimodal group. In ovarian cancer, it has allowed us to collaborate to explore new biomarkers[38,39] and develop new longitudinal algorithms.[28,40] The data provides a unique opportunity to study the natural history of ovarian and tubal cancer. Significant research is also underway on other cancers[41,42] and chronic conditions such as cardiovascular disease.[43] Details of access to this resource can be found at http://uklwc.mrcctu.ucl.ac.uk/ .

A key limitation is the interval from end of screening (2011) to censorship, which raises the possibility of dilution of the screening effect. However, extended follow-up

after screening is the norm in screening trials and did not impact on mortality reduction seen in the European Randomized Study of Screening for Prostate Cancer.[44]

The majority of women who were screen detected, were diagnosed and treated more than a decade ago. In retrospect, second line tests could have been further optimised so that time to diagnosis following an abnormal screen was reduced.[27] Clinicians could have been encouraged to intervene earlier when faced with rising biomarker levels and normal imaging.[45] Modelling suggests that the majority of Stage I Type II epithelial cancers are 0·4-1·3cm in diameter and therefore difficult to reliably image.[46] Finally, most of the screen detected women (2001-2011) did not have the advantage of more recent advances in clinical management (widespread use of ultraradical surgery, earlier treatment modulation based on better prognostic indicators, targeted therapies) that could have improved outcomes.

Conclusion

We began our quest to reduce deaths from ovarian cancer in the eighties,[5,6] started our pilot RCT in mid-nineties[47] and then undertook UKCTOCS, over the next 20 years. The journey has involved over 200000 women who have trusted us and given generously of their time, numerous NHS staff both at the trial centres and more widely and the support of many UK and international charities, funding agencies and expert groups. To them, we are hugely grateful. While disappointing, the trial has provided a clear answer that our screening strategies coupled with treatment protocols available in 2001-2011(the active screening phase) did not save lives. The important findings of a shift in stage distribution and the impact on stage specific disease mortality in the multimodal arm requires further analyses which will be reported in a separate publication. Currently general population screening for ovarian/tubal cancer cannot be recommended. We remain optimistic that further research will develop more effective ways to detect and treat this lethal disease. Meanwhile, the UKCTOCS biorepository with longitudinal samples provides a unique opportunity to advance early detection biomarker research.

Contributors

UM is the chief investigator (CI) and was Co-CI from 2001-2014. MP is the trial statistician. IJJ was CI from 2000-2014 and is a co-investigator as are SJS, SC, AM and LF. UM, MP, IJJ, SJS, SC, AM, LF, AGM, MB, AR and JKK contributed to the study concept and design. AGM, MB and UM performed the literature search. Data collection was performed by AR, AGM, CK, GC, JT, SM. Outcomes review was done by NS (chair), RM, RW, RA, AS, KW, LC. AR and UM prepared the dataset that was used for the analysis. MB, AR, SJS, YL, MR, AM undertook the statistical analysis under the supervision of MP. AR, MB, AGM, UM prepared the figures and tables and drafted the manuscript. All contributed to the interpretation of the data and revision of the manuscript. All authors approved the report before submission.

Acknowledgments

We thank the volunteers without whom the trial would not have been possible and everyone involved in conduct and oversight of UKCTOCS. We are very grateful to the current members of the UKCTOCS Trial Steering Committee - Prof Henry Kitchener (Chair), Prof Julietta Patnick, Prof Jack Cuzick and Ms Annwen Jones. We are indebted to the administrative support provided by Mrs Anna Widdup, Ms Roxanne Payne and Ms Jasvinder Dinza. We are indebted to the funding agencies for working together to support this long and challenging trial.

Funding

We are grateful to the funding agencies for their sustained support for UKCTOCS over many years. The Long Term Follow Up (LFTU) UKCTOCS is supported by National Institute for Health Research (NIHR HTA grant 16/46/01), Cancer Research UK (CRUK) and The Eve Appeal. UKCTOCS was funded by Medical Research Council (G9901012 and G0801228), CRUK (C1479/A2884), and the Department of Health, with additional support from The Eve Appeal. Researchers at UCL are supported by the NIHR University College London Hospitals (UCLH) Biomedical Research Centre and MRC CTU at UCL core funding (MR_UU_12023).

Disclaimer

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Research in Context

Evidence before the study

We searched PubMed between 2015 and 2020 using no language restrictions to find any published randomised control trials for ovarian cancer screening which reported mortality data. The following keywords were used to search the database: "ovarian cancer" AND "randomised controlled trial" AND "screening" AND "mortality". There were two relevant publications.

In UKCTOCS, at a median follow up of 11·1 years, there was no significant reduction in deaths from ovarian cancer in either of the screen groups (multimodal or ultrasound) compared to the ‘no screening’ group using Cox’s analysis. The reduction in deaths was delayed and only apparent after about seven years. There was a suggestion that 15% fewer women in the multimodal and 11% fewer in the ultrasound group died from ovarian cancer compared to the ‘no screening’ group. Additionally, a significantly greater proportion (13%) of women with ovarian cancers in the multimodal group but not in the ultrasound group were found at an earlier stage (Stage I and II) compared to the ‘no screening’ group. As the data did not definitively answer the question whether screening saved lives, follow-up was continued to gather more evidence.

The Ovarian Cancer Screening arm of the PLCO trial in the US is the only other large randomised controlled trial (n=78 216) to explore mortality benefit. Following extended follow-up (median 14·7 years), the trial confirmed previous findings of no ovarian cancer mortality reduction between the screen and control arms.

Added value of this study

Long term follow up (on average, more than 16 years following recruitment) in the largest ovarian cancer screening trial to date provides definitive new evidence that neither screening approaches used in UKCTOCS reduced deaths from ovarian cancer, compared to no screening. This was despite 47% increase in incidence of women with ovarian and tubal cancer diagnosed at stage I and 24·5% decrease in those diagnosed with stage IV in the multimodal group compared to the no screening group. Importantly, however, there was only a 10% decrease in overall incidence of Stage III/IV.

Implications of all the available evidence

General population screening for ovarian/tubal cancer with either approach as used in UKCTOCS cannot be recommended. We need a screening strategy that can detect ovarian and tubal cancer in asymptomatic women even earlier in its course and in a larger proportion of women than the tests used in the trial. Meanwhile our results emphasise the importance of having ovarian/tubal cancer mortality as the primary outcome in screening trials.

References

1. CRUK. Ovarian cancer survival statistics. 2018. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/survival#heading-Three (accessed 21/01/2021.

2. Jacobs IJ, Skates SJ, MacDonald N, et al. Screening for ovarian cancer: a pilot randomised controlled trial. *Lancet* 1999; **353**(9160): 1207-10.

3. van Nagell JR, Jr., Burgess BT, Miller RW, et al. Survival of Women With Type I and II Epithelial Ovarian Cancer Detected by Ultrasound Screening. *Obstet Gynecol* 2018; **132**(5): 1091-100.

4. Kobayashi H, Yamada Y, Sado T, et al. A randomized study of screening for ovarian cancer: a multicenter study in Japan. *Int J Gynecol Cancer* 2008; **18**(3): 414-20.

5. Campbell S, Bhan V, Royston P, Whitehead MI, Collins WP. Transabdominal ultrasound screening for early ovarian cancer. *BMJ* 1989; **299**(6712): 1363-7.

6. Jacobs I, Stabile I, Bridges J, et al. Multimodal approach to screening for ovarian cancer. *Lancet* 1988; **1**(8580): 268-71.

7. Buys SS, Partridge E, Black A, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 2011; **305**(22): 2295-303.

8. Pinsky PF, Yu K, Kramer BS, et al. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15years follow-up. *Gynecol Oncol* 2016; **143**(2): 270-5.

9. Jacobs IJ, Menon U, Ryan A, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* 2016; **387**(10022): 945-56.

10. Menon U, Gentry-Maharaj A, Ryan A, et al. Recruitment to multicentre trials--lessons from UKCTOCS: descriptive study. *BMJ* 2008; **337**: a2079.

11. Menon U, Gentry-Maharaj A, Hallett R, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol* 2009; **10**(4): 327-40.

12. UKCTOCS. Protocol for the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) and the Long Term Impact of Screening on Ovarian Cancer Mortality (LTFU UKCTOCS). 2020.

http://ukctocs.mrcctu.ucl.ac.uk/media/1066/ukctocs-protocol_v90_19feb2020.pdf (accessed 21/01/2021.

13. Skates SJ, Pauler DK, Jacobs IJ. Screening Based on the Risk of Cancer Calculation From Bayesian Hierarchical Changepoint and Mixture Models of Longitudinal Markers. *Journal of the American Statistical Association* 2001; **96**(454): 429-39.

14. Kalsi JK, Ryan A, Gentry-Maharaj A, et al. Completeness and accuracy of national cancer and death registration for outcome ascertainment in trials-an ovarian cancer exemplar. *Trials* 2021; **22**(1): 88.

15. Meinhold-Heerlein I, Fotopoulou C, Harter P, et al. The new WHO classification of ovarian, fallopian tube, and primary peritoneal cancer and its clinical implications. *Arch Gynecol Obstet* 2016; **293**(4): 695-700.

16. Daya D, Cheung AN, Khunamornpong S. Tumors of the peritoneum: epithelial tumors of Müllerian type. In: Kurman RJ CM, Herrington CS, Young RH, eds, ed. WHO classification of tumors of female reproductive organs. 4th ed. Lyon: International Agency for Research on Cancer; 2014: 92-3.

17. Tavassoli FAD, P.; eds. . Tumors of the breast and female genital organs. World Health Organization Classification of Tumours: Pathology and Genetics. Lyon: World Health Organization; 2003.

18. Burnell M, Gentry-Maharaj A, Skates S, et al. UKCTOCS Update: Applying insights of delayed effects in cancer screening trials to the long-term follow-up mortality analysis. *medRxiv* 2020.

19. Karrison TG. Versatile tests for comparing survival curves based on weighted log-rank statistics. *The Stata Journal* 2016; **16**(3): 678-90.

20. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002; **21**(15): 2175-97.

21. Li G, Taljaard M, Van den Heuvel ER, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol* 2017; **46**(2): 746-55.

22. Ahlbom A. Biostatistics for Epidemiologists. 1st Edition ed. Boca Raton, FL: CRC Press; 1993.

23. Rosenthal AN, Fraser LSM, Philpott S, et al. Evidence of Stage Shift in Women Diagnosed With Ovarian Cancer During Phase II of the United Kingdom Familial Ovarian Cancer Screening Study. *J Clin Oncol* 2017; **35**(13): 1411-20.

24. Skates SJ, Greene MH, Buys SS, et al. Early Detection of Ovarian Cancer using the Risk of Ovarian Cancer Algorithm with Frequent CA125 Testing in Women at Increased Familial Risk - Combined Results from Two Screening Trials. *Clin Cancer Res* 2017; **23**(14): 3628-37.

25. Gaba F, Goyal S, Marks D, et al. Surgical decision making in premenopausal BRCA carriers considering risk-reducing early salpingectomy or salpingo-oophorectomy: a qualitative study. *J Med Genet* 2021.

26. Dilley J, Allen P, Gentry-Maharaj A, et al. Symptoms reported by women with screen detected invasive epithelial ovarian cancer in the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *International Journal of Gynecological Cancer* 2015: 423-4.

27. Menon U, Ryan A, Kalsi J, et al. Risk Algorithm Using Serial Biomarker Measurements Doubles the Number of Screen-Detected Cancers Compared With a Single-Threshold Rule in the United Kingdom Collaborative Trial of Ovarian Cancer Screening. *J Clin Oncol* 2015; **33**(18): 2062-71.

28. Blyuss O, Burnell M, Ryan A, et al. Comparison of Longitudinal CA125 Algorithms as a First-Line Screen for Ovarian Cancer in the General Population. *Clin Cancer Res* 2018; **24**(19): 4726-33.

29. Tabar L, Yen AM, Wu WY, et al. Insights from the breast cancer screening trials: how screening affects the natural history of breast cancer and implications for evaluating service screening programs. *Breast J* 2015; **21**(1): 13-20.

30. Schoen RE, Pinsky PF, Weissfeld JL, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 2012; **366**(25): 2345-57.

31. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N Engl J Med* 2020; **382**(6): 503-13.

32. Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR. Lung cancer screening: the Mayo program. *J Occup Med* 1986; **28**(8): 746-50.

33. Kubik A, Parkin DM, Khlat M, Erban J, Polak J, Adamec M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int J Cancer* 1990; **45**(1): 26-33.

34. Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest* 1984; **86**(1): 44-53.

35. Doria-Rose VP, Marcus PM, Szabo E, Tockman MS, Melamed MR, Prorok PC. Randomized controlled trials of the efficacy of lung cancer screening by sputum cytology revisited: a combined mortality analysis from the Johns Hopkins Lung Project and the Memorial Sloan-Kettering Lung Study. *Cancer* 2009; **115**(21): 5007-17.

36. Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP, Fontana RS. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *J Natl Cancer Inst* 2006; **98**(11): 748-56.

37. NHSx. Accelerating Detection of Disease. https://www.nhsx.nhs.uk/ai-lab/ai-lab-programmes/accelerating-detection-of-disease/.

38. Russell MR, Walker MJ, Williamson AJ, et al. Protein Z: A putative novel biomarker for early detection of ovarian cancer. *Int J Cancer* 2016; **138**(12): 2984-92.

39. Whitwell HJ, Worthington J, Blyuss O, et al. Improved early detection of ovarian cancer using longitudinal multimarker models. *Br J Cancer* 2020; **122**(6): 847-56.

40. Gentry-Maharaj A, Blyuss O, Ryan A, et al. Multi-Marker Longitudinal Algorithms Incorporating HE4 and CA125 in Ovarian Cancer Screening of Postmenopausal Women. *Cancers (Basel)* 2020; **12**(7).

41. Widschwendter M, Evans I, Jones A, et al. Methylation patterns in serum DNA for early identification of disseminated breast cancer. *Genome Med* 2017; **9**(1): 115.

42. Jenkinson C, Elliott VL, Evans A, et al. Decreased Serum Thrombospondin-1 Levels in Pancreatic Cancer Patients Up to 24 Months Prior to Clinical Diagnosis: Association with Diabetes Mellitus. *Clin Cancer Res* 2016; **22**(7): 1734-43.

43. Borges MC, Schmidt AF, Jefferis B, et al. Circulating Fatty Acids and Risk of Coronary Heart Disease and Stroke: Individual Participant Data Meta-Analysis in Up to 16 126 Participants. *J Am Heart Assoc* 2020; **9**(5): e013131.

44. Schroder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* 2012; **366**(11): 981-90.

45. Menon UG-M, A.; Ryan, A., Burnell, M.; Dawnay, A,; Habib, M., Apostolidou, S.; Skates, S.; Parmar, M.; Jacobs, I. . Serial CA125 interpreted using the Risk of Ovarian Cancer algorithm can detect ovarian cancer in absence of ultrasound abnormalities (Oral presentation, Awarded 1st prize). London. Innovations & Progress

in Healthcare for Women, 2nd International Meeting: Prevention, Screening and Risk Prediction in Women's Health; 2009.

46. Brown PO, Palmer C. The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS Med* 2009; **6**(7): e1000114.

47. Menon U, Skates SJ, Lewis S, et al. Prospective study using the risk of ovarian cancer algorithm to screen for ovarian cancer. *J Clin Oncol* 2005; **23**(31): 7919-26.

Table and Figure legends

Table 1: Ovarian and tubal cancers grouped by primary site and screening status

Table 2: Summary of incidence and case fatality rate by stage (FIGO 2014) for ovarian and tubal cancer

Table 3: Summary of analyses of relative reduction of ovarian and tubal cancer deaths

Figure 1: Trial profile
*Footnote: MMS = multimodal screening. USS = ultrasound screening. *Events occurred before recruitment, but discovered after randomisation*

Figure 2: Kaplan-Meier cumulative mortality for ovarian and tubal cancer per 100 000 women
Footnote: MMS = multimodal screening. USS = ultrasound screening

Figure 3: Kaplan-Meier cumulative incidence for (A) all ovarian and tubal cancers and for (B) invasive epithelial ovarian and tubal cancers per 100 000 women
Footnote: MMS = multimodal screening. USS = ultrasound screening