Effect of organised cervical cancer screening on cervical cancer mortality in Europe: a systematic review

Erik E.L. Jansen a,*, Nadine Zielonke a, Andrea Gini a, Ahti Anttila b, Nereo Segnan c, Zoltán Voko d,e, Urška Ivanuš f, Martin McKee g, Harry J. de Koning a, Inge M.C.M. de Kok a on behalf of the EU-TOPIA consortium1

a Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
b Finnish Cancer Registry, Helsinki, Finland
c Department of Cancer Epidemiology, CPO Piemonte, City of Health and Science University Hospital, Turin, Italy
d Department of Health Policy & Health Economics, Eötvös Loránd University, Budapest, Hungary
e Syreon Research Institute, Budapest, Hungary
f Institute of Oncology Ljubljana, Ljubljana, Slovenia
g London School of Hygiene and Tropical Medicine, London, England, UK

Received 11 November 2019; accepted 2 December 2019
Available online 21 January 2020

KEYWORDS
Uterine cervical neoplasms; Early detection of cancer; Mortality; Europe; Systematic review

Abstract Background: Organised cervical cancer (CC) screening programmes are delivered in many different ways across the European Union and its regions. Our aim was to systematically review the impact of these programs on CC mortality.

Methods: Two independent reviewers identified all eligible studies investigating the effect of organised screening on CC mortality in Europe. Six databases including Embase, Medline and Web of Science were searched (March 2018) with predefined inclusion and exclusion criteria. Only original studies with at least five years of follow-up were considered. Validated tools were used to assess the risk of bias of the included studies.

Results: Ten observational studies were included: seven cohort and three case-control studies. No randomised controlled trials were found, and there were no eligible studies from the eastern and southern part of Europe. Among the eligible studies, seven were conducted in the twentieth century; they scored lower on the risk of bias assessment. CC mortality reduction for women attending organised screening vs. non-attenders ranged from 41% to 92% in seven studies. Reductions were similar in Western (45–92%) and Northern (41–87%) Europe and

* Corresponding author: Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Dr. Molewaterplein 40, Rotterdam, 3015 GD, the Netherlands.
E-mail address: e.e.l.jansen@erasmusmc.nl (E.E.L. Jansen).
1 Members in the collaboration group “EU-TOPIA consortium” are listed in appendix section.

0959-8049/ © 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Cervical screening has been shown to reduce the incidence and mortality of cervical cancer (CC). Precancerous lesions can be treated, preventing progression to invasive disease [1–6], thereby avoiding the need for chemotherapy or radiotherapy or infertility due to removal of the cervix [7–9]. Screening is estimated to reduce the incidence rate of CC by 50–60% [10]. Yet although most European countries have offered some sort of CC screening for decades [11], 34,000 new CC cases are detected in Europe each year, with 13,000 deaths [12].

CC screening is most effective where it is undertaken within an organised programme. Yet, so far, 19 of the 28 countries of the European Union have yet to implement such programmes despite clear recommendations agreed by the European Council [13–15]. These recommendations advocate starting screening at an age between 20 and 30, repeating at three to five year intervals until the age of 60 or 65 [16,17]. Screening can be performed using a Pap smear to detect any abnormal cells (cytology) and/or a test to check for the presence of the human papilloma virus (HPV), the causal agent in CC [17,18]. While ten countries are currently rolling out an organised CC screening programme and three countries are currently planning or piloting such programme, six countries only have a non-population—based programme or no programme at all [13,14].

The reduction in mortality that can be achieved by a screening programme depends on several factors. These include the epidemiology of HPV infection in the population and characteristics of the screening programme, including the starting and stopping ages, screening interval and coverage [19]. Other factors include the performance of screening activities, in terms of sensitivity and specificity, access to treatment by those in whom lesions are detected and quality of follow-up. These parameters can vary widely so it is likely that observed reductions in CC mortality will also vary. However, the extent to which screening does achieve reductions in Europe, including differences among countries and over time, has not previously been brought together systematically.

This systematic review is part of the EU-TOPIA (TOwards imProved screening for breast, cervical and colorectal cancer In All of Europe) project that is evaluating and quantifying the harms and benefits of cancer screening in European countries, to improve health outcomes and increase equity. Both women and policymakers should know whether their screening programmes are performing optimally and the scale of the benefit that they can expect. This review seeks to address this question by searching for the best quality published evidence on the effect of screening on reducing CC mortality in the European region. This can be used to benchmark progress in countries with existing screening programmes and those that are starting new ones.

2. Methods

The systematic review was part of a large one on cancer mortality reduction associated with screening for breast, colorectal and CC within the EU-TOPIA project. The protocol of this systematic review was published in PROSPERO (International Prospective Register of Systematic Reviews, CRD42016042433) on July 6, 2016 [20]. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement served as a guideline for the performance of this systematic review [21].

2.1. Search strategy

The PICOS (population, intervention, control, outcome and study design) criteria (Table 1) from the study protocol [20] formed the basis of the search strategy in six electronic databases. Embase, Ovid Medline, Web of Science, PubMed, Google Scholar and the Cochrane Library were searched from inception until March 2018 for articles in English related to population-based CC screening and effects on CC mortality in European countries using a computer-assisted search code compiled by a research librarian (appendix Table A). Experts from the field were invited to suggest additional relevant articles and grey literature to be added to the list of potentially eligible articles. We also manually searched reference lists of pertinent articles to find any relevant citations that our searches might have missed. Duplicates were removed, and the remaining references were managed using Thomson Reuters Endnote X7.5.

2.2. Study selection

All retrieved references were screened for title and abstract by two investigators (E.J. and N.Z.)
studies was also extracted. All relevant outcome measures are presented in a table, sorted by European region as defined by the EuroVoc multilingual thesaurus of the European Union [24].

When comparing reductions in CC mortality across studies, odds ratios were interpreted as being RRs as these will be very similar as long as the incidence of the event is less than 10% in the overall population [25]. This also applies to case-control studies [26].

To assess the risk of bias of included studies, two investigators independently scored the risk of bias using the Newcastle-Ottawa Scale (NOS) for observational studies, covering the domains of study group selection, comparability between the study groups, outcome measurement and exposure to the intervention [27]. Because of the differences in the study design between cohort studies and case-control studies, different types of biases could occur so the NOS applies different questions to each [27]. A higher score on the NOS corresponds to less risk of bias. The NOS initially did not award a point for adequate case definition of case-control studies if this was based on record linkage only. Anttila et al. [28] advocated that cancer registries, when mandated and resourced, should take co-responsibility in the evaluation of the quality and impact of organised screening. In addition, as cancer registries in some countries have a very high percentage of histologically verified cases, which we designated as independent validation, we did award a point for this if the percentage was known to be more than 95% according to the International Agency for Research on Cancer [29]. All scoring discrepancies between reviewers were discussed until consensus was reached. Final decisions about remaining discrepancies were made by the third investigator. Studies were sorted by NOS score in the result tables providing the opportunity for readers to interpret the results accordingly.

3. Results

The number of references remaining at each step of the study selection process is shown in the PRISMA flow chart in Fig. 1. The initial search of the six
databases resulted in 2562 records. Experts from the field added 61 references, while the reference lists identified a further 15 studies to be screened, adding up to 76 extra references. After removing duplicates, 1816 studies remained for title/abstract screening. The two independent reviewers reached consensus to include 66 articles that potentially fitted the predefined PICOS criteria for full-text review. Two studies were excluded because the full-text articles appeared to be unavailable in English. Of the remaining 64 articles, 54 were excluded for a variety of other reasons (Fig. 1). Ultimately, this review included a total of ten studies. All included studies were present in the initial database search. Excluded articles are listed in appendix Table B including reasons for exclusion.

3.1. Characteristics of the included studies

All included studies were observational, and no randomised controlled trials (RCTs) were found. Of the ten studies included, seven were performed in Northern European countries and three were in Western European countries (Table 2). Three case-control studies were included, whereas the other references were all from cohort studies. The case-control studies included 108, 198 and 110,619 cases, whereas the control groups comprised 216, 1218 and ~23,000 subjects, respectively [30–32]. The sample sizes in the cohort studies varied from 15,257 to about 4,200,000 women, although not all reported the exact sample size. Across studies, the starting age for screening was between 20 and 35 while screening was performed until age 49 to 69 with intervals ranging between two and five years. One study [31] did not report the screening interval. The follow-up time was at least five years in the cohort studies [33], with a maximum of 36 years [34]. Adherence in the cohort studies ranged from 72% to 86%, although this information was missing from three studies [33–35]. The year in which the studies were published ranged from 1979 to 2016 with most published before 1995.

3.2. Risk of bias

The risk of bias varied between studies (Table 2). Three studies [31,33,36] scored four of nine points on the NOS, four studies scored five or six points and three studies [30,32,37] scored seven, eight or nine points. Two of three studies from Western Europe were among those
Table 2
Characteristics, risk of bias and results on cervical cancer mortality of included studies, by region (based on EuroVoc) and quality score.

<table>
<thead>
<tr>
<th>Region/Study</th>
<th>Country</th>
<th>Study type</th>
<th>Participants</th>
<th>Target age (years)</th>
<th>Screening interval (years)</th>
<th>Follow-up (years)</th>
<th>Adherence (%)</th>
<th>Correction for self-selection bias</th>
<th>NOS score</th>
<th>RR (95% CI) for cervical cancer mortality (invited)</th>
<th>RR (95% CI) for cervical cancer mortality (participated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Europe [24]</td>
<td>Lönberg S, 2013 [30]</td>
<td>Finland</td>
<td>Case control</td>
<td>198 cases (71(^b)) 1218 controls(^c) (876(^b))</td>
<td>25–69</td>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
<td>9/9</td>
<td>0.34 (0.14–0.49)</td>
</tr>
<tr>
<td></td>
<td>Dugue P, 2014 [37]</td>
<td>Denmark</td>
<td>Cohort</td>
<td>903,439 participants 253,232 non-participants(^c)</td>
<td>23–59</td>
<td>3</td>
<td>13</td>
<td>78</td>
<td>No</td>
<td>7/9</td>
<td>0.13 (0.11–0.15)</td>
</tr>
<tr>
<td></td>
<td>Bergström S, 1999 [34]</td>
<td>Sweden</td>
<td>Cohort</td>
<td>24,389 cases 10,655 deaths</td>
<td>30–49</td>
<td>4</td>
<td>26–36</td>
<td>86(^d)</td>
<td>Invited vs. non-invited Invited vs. non-invited</td>
<td>6/9</td>
<td>0.21 (N/A)</td>
</tr>
<tr>
<td></td>
<td>Mählick C, 1994 [35]</td>
<td>Sweden</td>
<td>Cohort</td>
<td>~4,200,000 invited</td>
<td>30–49</td>
<td>4–5</td>
<td>22</td>
<td>86(^d)</td>
<td>Invited vs. non-invited Yes</td>
<td>6/9</td>
<td>0.47 (0.28–77)</td>
</tr>
<tr>
<td></td>
<td>Berget A, 1979 [48]</td>
<td>Denmark</td>
<td>Cohort</td>
<td>13,148 participants 2109 non-participants(^c)</td>
<td>30–49</td>
<td>4–5</td>
<td>6–8</td>
<td>86</td>
<td>Invited vs. non-invited Yes</td>
<td>5/9</td>
<td>0.70 (N/A)(^e) 0.16 (N/A)</td>
</tr>
<tr>
<td></td>
<td>Lynge E, 1989 [49]</td>
<td>Denmark</td>
<td>Cohort</td>
<td>N/A</td>
<td>20–59</td>
<td>2</td>
<td>15</td>
<td>72–81</td>
<td>Invited vs. non-invited Yes</td>
<td>4/9</td>
<td>0.68 (0.59–0.78)</td>
</tr>
<tr>
<td></td>
<td>Magnus K, 1987 [36]</td>
<td>Norway</td>
<td>Cohort</td>
<td>45,960 invited</td>
<td>25–59</td>
<td>2–4</td>
<td>24</td>
<td>76</td>
<td>Yes</td>
<td>4/9</td>
<td>0.83 (N/A) 0.59 (N/A)</td>
</tr>
<tr>
<td>Western Europe [24]</td>
<td>Landy R, 2016 [32]</td>
<td>United Kingdom</td>
<td>Case control</td>
<td>11,619 cases ~23,000 controls(^c)</td>
<td>35–64</td>
<td>3–5</td>
<td>5</td>
<td>N/A</td>
<td>No</td>
<td>8/9</td>
<td>0.08 (0.07–0.09)</td>
</tr>
<tr>
<td></td>
<td>Macgregor E, 1994 [31]</td>
<td>Scotland</td>
<td>Case control</td>
<td>108 cases (38(^b)) 216 controls(^c) (157(^b))</td>
<td>25–60</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>4/9</td>
<td>0.25 (0.12–0.48)</td>
</tr>
<tr>
<td></td>
<td>Ebeling K, 1986 [33]</td>
<td>Germany</td>
<td>Cohort</td>
<td>N/A</td>
<td>20–64</td>
<td>2</td>
<td>5</td>
<td>N/A</td>
<td>No</td>
<td>4/9</td>
<td>0.09 (N/A)</td>
</tr>
</tbody>
</table>

NOS, Newcastle-Ottawa Scale; CI, confidence interval; N/A, not available; RR, relative risk (odds ratio and percent reduction are presented as a RR because they are similar because of the relative low incidence of cervical cancer) \[25\].

Target age: Ages targeted by the organized screening programme; Follow-up: Follow-up time after initiation of the screening programme.

\(^a\) Quality assessment made according to the Newcastle-Ottawa Scale.

\(^b\) Exposed to screening.

\(^c\) Controls were drawn from the same population as the intervention group.

\(^d\) Estimation based on Stenkvist et al. (1984), women with 2 or more smears over a ten-year period \[47\].

\(^e\) Not significant.
with four points. The most common criteria that affected the risk of bias assessment were a lack of information on the prevalence of CC at the start of the study and on the length of the follow-up. Furthermore, authors often failed to correct for any differences between the intervention and control groups (e.g. age) and did not specify the method by which the exposure to the intervention was measured. Tables C.1 and C.2 in the appendix provide the arguments for each score on the NOS. An overview of all conflict of interest statements is provided in appendix Table D.

3.3. CC mortality outcomes

All included studies reported a reduction in CC mortality (Table 2) in those attending screening compared with non-attenders (Fig. 2a) and in those invited for screening compared with non-invited women (Fig. 2b), although not all studies reported whether this reduction was found to be significant or not. In cohort studies reporting the effect of inviting a population for screening, the CC mortality reduction was between 17% and 79% in Northern Europe; no such studies were performed in other regions. In cohort studies from Northern Europe, CC mortality reduction among those participating in screening was between 41% and 84% in studies that corrected for self-selection bias and 87% in the study that did not do so. The cohort study from Western Europe did not correct for self-selection bias and reported a CC mortality reduction of 91%. In the case-control study in Northern Europe, the OR of dying from CC after participating in screening compared with non-participants, and corrected for self-selection bias, was 0.34 (95% CI: 0.14–0.49) [30]. In the two case-control studies from Western Europe, the ORs were 0.08 (95% CI: 0.07–0.09) [32] and 0.25 (95% CI: 0.12–0.48) [31], both uncorrected for self-selection bias.

4. Discussion

This systematic review shows that there is relatively little published evidence on CC mortality reduction after implementation of organised CC screening in Europe. No RCTs were reported from Europe, and there were no studies of any sort from Southern and Eastern Europe. However, in Northern Europe, two recent high-quality observational studies were reported. One showed a 51–86% CC mortality reduction after participation in organised screening, corrected for self-selection bias [30], and the other showed a 85–89% CC mortality reduction without correction for self-selection bias. In the 1990s, two large cohort studies were performed, showing a 23–72% CC mortality reduction after invitation for screening, although the target age range in the screening programmes at that time was still restricted to 30–49 years. In Western Europe, one recent high-quality study was conducted showing a 92% CC mortality reduction without correction for self-selection bias. The other two studies from Western Europe were both conducted before 1995 and scored four points on the NOS scale. Overall, therefore, the evidence confirms that screening for CC is associated with CC mortality reduction.

This conclusion is in line with a previous review performed by Peirson et al. [38] who found an association between screening for CC and CC mortality reduction. However, Peirson et al. [38] mainly found incidence reduction studies during their search, while focussing on a single Indian RCT when assessing CC mortality reduction (RR = 0.65 [95% CI = 0.47, 0.90]). The conclusions of this trial might not be applicable to the European setting as the background risk is likely to be different and the trial assessed the effect of a single lifetime screening only. To our knowledge, no other systematic reviews have been performed on the effect of CC screening on CC mortality.

The main strength of this systematic review is that it summarises all current existing evidence on the effects of CC screening on CC mortality in Europe while also providing information on the risks of bias within those studies. The scale of the reduction in CC mortality that we found is an important reason to implement or improve a CC screening programme. The effect size of the included studies varied, which can be explained by various reasons. Differences in CC mortality reduction between an invited and an attending population will depend on attendance rates of the invited population. Within the group of attenders, differences in CC mortality reduction can be expected based on follow-up time, general background risk, demographics of the study population and the characteristics of the screening programme and screening tests [19]. Furthermore, differences in effect size can be caused by selection of any control group if there is no correction applied for any differences in background risk, and the effect size can also be influenced by differences in treatment effectiveness over time and between countries [1]. To quantify the effects of each of the parameters affecting CC mortality reduction, simulation models could be used. These simulation models can mimic the natural history of CC in a population and apply different screening scenarios to compare expected CC mortality reduction between those scenarios. Despite all the different study settings in the studies found in this systematic review, there was a clear reduction in CC mortality in all ten included studies, of which most showed a statistical significant difference. This provides robust evidence that organised CC screening is able to reduce CC mortality.

Although all included studies found reductions in CC mortality associated with CC screening, some limitations must be mentioned. First, grey literature and non-English literature were excluded. Second, the NOS used to assess the risk of bias in the eligible studies was not specifically created for assessing screening intervention
Fig. 2. (a) Reduction in cervical cancer mortality after attending screening by European region [24] and NOS score. NOS = Newcastle-Ottawa scale (i.e. a higher score is a lower risk of bias) [27]; # Uncorrected for self-selection bias; Confidence intervals are shown as error bars if they were reported in the corresponding study. Fig. 2. (b) Reduction in cervical cancer mortality after being invited for screening in Northern Europe [24] by NOS score. NOS = Newcastle-Ottawa scale (i.e. a higher score is a lower risk of bias) [27]; Confidence intervals are shown as error bars if they were reported in the corresponding study. *Estimation based on Stenkvist et al. [47], women with 2 or more smears over a ten-year period.
Intraepithelial neoplasia or cancer in trials [43] and specificity than cytology in identifying cervical primary HPV test. HPV tests achieve higher sensitivity to other screening programmes, for instance, using a primary test, the conclusions cannot be directly applied because all of the included studies used cytology as a base. The conclusions will depend on the characteristics of the screening test.

Because all of the included studies used cytology as a primary test, the conclusions cannot be directly applied to other screening programmes, for instance, using a primary HPV test. HPV tests achieve higher sensitivity and specificity than cytology in identifying cervical intraepithelial neoplasia or cancer in trials [43–45]. Thus, it is reasonable to assume that CC mortality reduction with programmes using a primary HPV test will be at least as high as the cytology-based programmes. In 2017, the Dutch-organised screening programme was the first in the world to use the HPV test as a primary screening test. Monitoring of the outcomes of this and similar programmes will provide important information on the effect of HPV screening in practice.

This systematic review focusses on the main goal of screening: reduction of CC mortality. Yet, despite all its benefits, screening can also cause harms. Healthy women who undergo screening can be anxious about the outcome, overdiagnosed or overtreated [46]. Thus, when deciding on the optimal screening strategy for a specific country, one should always weigh the benefits against the harms, a consideration that was outside the scope of this systematic review.

Our main conclusion is that even though organised screening programmes have been running for many years, there is still relatively little known about the effect of screening on CC mortality reduction. However, studies from Finland, Sweden and the United Kingdom show significant and large effects, confirming the view that organised CC screening can reduce CC mortality. These results could be used as a benchmark for other European countries using similar methods. In the absence of evidence for a specific programme or country, modelling could be used to quantify the effects of individual characteristics of screening programs and the population on CC mortality reduction for each country.

Funding

This work was supported by the EU-Framework Programme (Horizon 2020) of the European Commission [project reference 634753; PI: prof HJ de Koning, MD PhD, Erasmus MC] that funded the EU-TOPIA project of which this systematic review is part of.

Role of the funding source

The funders had no influence on the outcomes of this systematic review.

Conflict of interest

None declared.

Acknowledgements

The authors would like to thank Wichor Bramer, biomedical information specialist at the medical library of Erasmus Medical Centre (Rotterdam, the Netherlands) for his valuable help, expert inputs and solutions in the making of this systematic review.
Appendices

Collaborators

Piret Veerus
Ahti Anttila
Sirpa Heinävaara
Tytti Sarkeala
Marcell Csándi
Janos Pitter
György Szél
Zoltán Vokó
Silvia Minozzi
Nereo Segnan
Carlo Senore

Marjolein van Ballegooijen
Inge Driessprong - de Kok
Andrea Gini
Eveline Heijnsdijk
Harry de Koning
Iris Lansdorp — Vogelaar
Nicolien van Ravesteyn
Nadine Zienonke
Urska Ivanuš
Katja Jarm
Dominika Novak Mlakar
Maja Primic-Zaželj
Martin McKee
Jennifer Priaulx

Appendix Table A
Computer-assisted search code according to reference databases.

<table>
<thead>
<tr>
<th>Source</th>
<th>Selection code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Embase</strong></td>
<td>('uterine cervix tumor'/exp OR 'Papanicolaou test'/de OR 'uterine cervix cytology'/de OR (((cervix* OR cervical*) NEAR/10 (cancer* OR neoplas* OR tumor* OR carcino* OR adenoacarcin* OR cytolog*)) OR Papanicolaou OR (pap NEXT/1 (smear* OR stain* OR test*)) OR (vagina* NEAR/3 smear*)):ab,ti) AND (screening/exp OR 'early diagnosis'/de OR (screen* OR (familial* OR periodic*) NEAR/3 examination*)):ab,ti) NOT (letter OR cousin OR editor OR congresses OR abstracts).pt. AND (english OR french OR german OR italian OR dutch OR hungarian OR polish OR belarus OR russia OR serbia OR slovakia OR slovenia OR ukraine OR france OR belgium OR albania OR italian OR irish OR english OR french OR german OR russian OR polish OR dutch OR french OR italian OR swedish OR spanish OR german OR dutch OR belgium OR france OR english OR german OR irish OR italian OR dutch).ab,ti) AND ('observational study'/exp OR 'cohort analysis'/exp OR 'longitudinal study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'health survey'/de OR 'health care survey'/de OR 'epidemiological data'/de OR 'case control study'/de OR 'cross-sectional study'/de OR 'correlational study'/de OR 'population research'/de OR 'family study'/de OR 'major clinical study'/de OR 'multicenter study'/de OR 'comparative study'/de OR 'follow up'/de OR 'clinical study'/de OR 'clinical article'/de OR 'clinical trial'/exp OR 'randomization'/exp OR 'intervention study'/de OR 'open study'/de OR 'community trial'/de OR 'review'/exp OR 'systematic review'/exp OR (organ* OR epidemiol* OR familial OR comparativ* OR communit*) NEAR/6 (stud* OR data OR research) OR (cohort* OR longitudinal* OR retrospective* OR prospective* OR population* OR (national* NEAR/3 (stud* OR survey)) OR (health* NEAR/3 survey*) OR (case OR cases OR match*)) NEAR/3 (control*) OR (next NEXT/1 section*) OR correlation* OR multistorey* OR (multi* NEXT/1 center*) OR 'follow up' OR followup* OR clinical* OR trial OR random* OR review*):ab,ti)</td>
</tr>
<tr>
<td><strong>Ovid Medline</strong></td>
<td>(&quot;Uterine Cervical Neoplasms&quot;/OR &quot;Papanicolaou test&quot;/OR &quot;Vaginal Smears&quot;/OR (((cervix* OR cervical*) ADJ10 (cancer* OR neoplas* OR tumor* OR carcino* OR adenoacarcin* OR cytolog*)):ab,ti) AND (screening/exp OR 'early Diagnosis'/exp OR (screen* OR (familial* OR periodic*) ADJ3 examination*)):ab,ti) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt. AND (english OR french OR german OR italian OR dutch).ab,ti) AND ('observational study'/exp OR 'cohort studies'/exp OR 'Health Surveys'/exp OR 'Epidemiologic Studies'/exp OR 'Case-Control Studies'/exp OR 'Cross-Sectional Studies'/exp OR 'multicenter study'/exp OR 'comparative study'/exp OR 'clinical study'/exp OR 'clinical trials'/exp OR 'Random Allocation'/exp OR 'review'/exp (organ* OR epidemiol* OR familial OR comparativ* OR communit*) NEAR/6 (stud* OR data OR research) OR (cohort* OR longitudinal* OR retrospective* OR prospective* OR population* OR (national* NEAR/3 (stud* OR survey)) OR (health* NEAR/3 survey*) OR (case OR cases OR match*) NEAR/3 (control*) OR (next NEXT/1 section*) OR correlation* OR multistorey* OR (multi* NEXT/1 center*) OR 'follow up' OR followup* OR clinical* OR trial OR random* OR review*):ab,ti)</td>
</tr>
</tbody>
</table>

(continued on next page)
Appendix Table A (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Selection code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web of science</td>
<td>TS=((((cervix* OR cervical*) NEAR/10 (cancer* OR neoplas* OR tumor* OR carcino* OR adenocarcin* OR cytology*) OR Papanicolaou OR (pap NEXT/1 (smear* OR stain* OR test*)) OR (vagina* NEAR/3 smear*))))ab,ti) AND (screen* OR ((annual* OR periodic*) NEAR/3 examination*)) OR (early NEAR/3 (diagnos* OR detect*))) AND ((mortalit* OR (death NEAR/1 rate*)):ab,ti) AND (mortalit* OR (death NEXT/1 rate*)):ab,ti) AND ((europe* OR Andorra* OR Austria* OR Balkan* OR Belgi* OR Albania* OR Baltic-State* OR Bosnia* OR Herzegovina* OR Bulgaria* OR Croatia* OR Czech* OR Hungarian* OR Kosovo* OR Macedonia* OR Moldova* OR Monteneg* OR Poland* OR polish* OR Belarus* OR Romania* OR Russia* OR Serbia* OR Slovak* OR Slovenia* OR Ukraine* OR France* OR french* OR German* OR Gibraltar* OR Great-Brit* OR uk OR united-kingdom* OR England* OR Scotland* OR Wales* OR welsh OR Greece* OR Ireland* OR Italy OR Italian OR Liechtenstein* OR Luxembourg* OR Monaco* OR Netherlands* OR dutch OR holland OR Portugal* OR San-Marino* OR Scandinavia* OR Nordic* OR Denmark* OR danish OR Finland* OR finnish OR Ireland* OR Norwa* OR norwegian OR Sweden* OR swedish OR Spain* OR spanish OR Switzerland* OR swiss)).</td>
</tr>
<tr>
<td>PubMed publisher</td>
<td>((observation* OR epidemiolog*) ADJ6 (stud* OR data OR research)) OR cohort* OR longitudinal* OR retrospective* OR prospectiv* OR population* OR (national* ADJ3 (stud* OR survey)) OR (health* ADJ3 survey*) OR ((case OR cases OR match*) ADJ3 control*) OR (cross ADJ1 section*) OR correlation* OR multicenter* OR (multi* ADJ center*) OR &quot;follow up&quot; OR followup* OR clinical* OR trial OR random* OR review*).ab,ti.</td>
</tr>
</tbody>
</table>

Appendix Table B
Characteristics of excluded studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Aa M, 1993</td>
<td>The study did not provide absolute numbers of events and participants or a relative risk.</td>
<td>Kostova P, 2010</td>
<td>No group sizes reported and no relative risk.</td>
</tr>
<tr>
<td>Aareleid T, 1993</td>
<td>Rates before and after screening implementation. Causal relation between screening and mortality reduction not tested.</td>
<td>Kovács A, 2008</td>
<td>Study provides no mortality or incidence data.</td>
</tr>
<tr>
<td>Adami H, 1994</td>
<td>The study did not provide absolute numbers of events or a relative risk.</td>
<td>Laara E, 1987</td>
<td>The study did not provide absolute numbers of events and participants or a relative risk.</td>
</tr>
<tr>
<td>D’Aló D, 2010</td>
<td>Causal relation between screening and mortality reduction not tested.</td>
<td>Lönnberg S, 2012</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Anttila A, 1999</td>
<td>Control group receives screening.</td>
<td>Mäjek O, 2016</td>
<td>Follow-up &lt;5 years.</td>
</tr>
<tr>
<td>Anttila A, 2007</td>
<td>Provides an overview of other studies.</td>
<td>Minelli L, 2007</td>
<td>The study did not provide absolute numbers of events and participants or a relative risk.</td>
</tr>
<tr>
<td>Anttila A, 2011</td>
<td>Control group receives screening.</td>
<td>Murphy M, 1987</td>
<td>The study did not provide absolute numbers of events and participants or a relative risk.</td>
</tr>
<tr>
<td>Apostol I, 2010</td>
<td>No baseline measurement.</td>
<td>Nieminen P, 1995</td>
<td>No number of events reported and no relative risk.</td>
</tr>
<tr>
<td>Arbyn M, 2012</td>
<td>Provides an overview of other studies.</td>
<td>Nieminen P, 1999</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Bojar I, 2012</td>
<td>Presented results are on absolute reductions.</td>
<td>Nowakowski A, 2012</td>
<td>No number of events reported and no relative risk.</td>
</tr>
<tr>
<td>Castillo M, 2018</td>
<td>Full article not available in English.</td>
<td>Nygard J, 2002</td>
<td>No absolute number of events. Control group receives screening.</td>
</tr>
<tr>
<td>Comber H, 2004</td>
<td>Causal relation between screening and mortality reduction not tested.</td>
<td>Parazzini F, 1990</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Crocetti E, 2007</td>
<td>No mortality outcomes, only incidence</td>
<td>Petterson F, 1995</td>
<td>No number of events reported and no relative risk.</td>
</tr>
<tr>
<td>Day N,</td>
<td>Provides an overview of other studies.</td>
<td>Quinn M, 1999</td>
<td>No number of events reported and no relative risk.</td>
</tr>
<tr>
<td>Ferraroni M, 1989</td>
<td>No mortality or incidence data provided.</td>
<td>Ronco G, 2005</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Gad C, 1976</td>
<td>Causal relation between screening and mortality reduction not tested.</td>
<td>Sasieni P, 2009</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Hakama M, 1976</td>
<td>Intervention group was after first negative smear.</td>
<td>Serraino D, 2015</td>
<td>No mortality outcomes, only incidence</td>
</tr>
<tr>
<td>Johannesson G, 1982</td>
<td>Presented results do not take screening history into account. No relative risk.</td>
<td>Simonella L, 2013</td>
<td>No number of cases reported. No control group.</td>
</tr>
<tr>
<td>Karzmarek-Borowska B, 2013</td>
<td>Full article not available in English.</td>
<td>Timonen S, 1974</td>
<td>Letter, No number of study groups or a relative risk.</td>
</tr>
<tr>
<td>Kinney W, 2003</td>
<td>Study not performed in Europe (U.S.A.)</td>
<td>Timonen S, 1974</td>
<td>No number of study groups or a relative risk.</td>
</tr>
<tr>
<td>Kok I, 2011</td>
<td>Causal relation between screening and mortality reduction not tested.</td>
<td>Timonen S, 1977</td>
<td>The study did not provide absolute numbers of events and participants.</td>
</tr>
<tr>
<td>Study</td>
<td>Case definition</td>
<td>Representativeness of the cases</td>
<td>Control selection</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------------------------</td>
<td>--------------------------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Landy R, 2016</td>
<td>Yes, but no actual deaths were counted, calculated number of deaths based on survival data were used.</td>
<td>Consecutive series of cases. All women who had cervical cancer between April 2007 and March 2013. (*)</td>
<td>Community controls. “All women registered with an NHS GP who did not have cervical cancer at the time of diagnosis were eligible as a control.” (*)</td>
</tr>
<tr>
<td>Lönnberg S, 2013</td>
<td>Record linkage with the cancer register. Cancer register has 99.5% histologically verified cases [29]. (*)</td>
<td>Consecutive series of cases. All registered cervical cancer deaths from the years 2000–2009. (*)</td>
<td>Community controls, drawn from the population register. (*)</td>
</tr>
<tr>
<td>Macgregor E, 1994</td>
<td>Yes, the records of the cases were obtained from hospital records and the cytopathology database. (*)</td>
<td>Not stated.</td>
<td>No description.</td>
</tr>
</tbody>
</table>

(*) The presence of this symbol means the study fitted the selected criteria and it was accounted in the final result. NHS, National Health Service; GP, general practitioner.
<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Selection of the non-exposed</th>
<th>Ascertainment of exposure</th>
<th>Absence of interest outcome at start of study</th>
<th>Comparability</th>
<th>Any additional factors</th>
<th>Outcome</th>
<th>Follow-up Length (&gt;8 years)</th>
<th>Adequacy of follow-up</th>
<th>Final result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berget A, 1979</td>
<td>Truly representative of the average Danish female population between 30 and 49 years old as all were invited for screening. (*)</td>
<td>Drawn from the same community as the exposed cohort. Not attending women. (*)</td>
<td>Invitation lists and meeting lists were used to measure invitation and attendance. (*)</td>
<td>No. Women with prior cervical lesion or hysterectomy were excluded. (*)</td>
<td>Yes. No. Quote: “The size of the study population does not allow correction for this.”</td>
<td>No.</td>
<td>Record linkage using death certificates and data from the Danish Cancer Registry. (*)</td>
<td>No. 6–8 years.</td>
<td>Yes. Only 0.64% were not identified and excluded. Emigrated women were followed and well described. (*)</td>
<td>6/9</td>
</tr>
<tr>
<td>Bergström S, 1999</td>
<td>Truly representative of the average Swedish female population between 30 and 49 years old as all were invited for screening. (*)</td>
<td>Drawn from the same community as the exposed cohort. Same population in the years before implementation of screening. (*)</td>
<td>Amount of smears reported, but the method of ascertainment is not described. (*)</td>
<td>No.</td>
<td>Standardised for age to the Swedish census population in 1970. (*)</td>
<td>Study corrected for period and cohort. (*)</td>
<td>Record linkage using the Swedish Cancer Register and the Swedish Cause of Death Register at Statistics Sweden. (*)</td>
<td>Yes. 26–36 years.</td>
<td>No statement.</td>
<td>6/9</td>
</tr>
<tr>
<td>Dugue P, 2014</td>
<td>Truly representative of the average Danish population between 23 and 51 years old, as all had the possibility to participate in two rounds of screening. (*)</td>
<td>Drawn from the same community as the exposed cohort. Not attending women. (*)</td>
<td>Secure record. Data on samples were retrieved from the Danish Pathology Data Bank, the National Health Service Register and the National Patient Register. (*)</td>
<td>No.</td>
<td>Yes. (*)</td>
<td>No.</td>
<td>Record linkage between the Danish Civil Registration System and the Danish Cause of Death Register using personal ID numbers. (*)</td>
<td>Yes. 13 years.</td>
<td>Small number lost. Unlikely to introduce bias. (*)</td>
<td>6/9</td>
</tr>
<tr>
<td>Ebeling K, 1986</td>
<td>Somewhat representative of the average Berlin women between 20 and 64 years old. (*)</td>
<td>Drawn from the same community as the exposed cohort. Not attending women. (*)</td>
<td>No description of method. Screening histories of all patients with cervical cancer were carefully monitored.</td>
<td>No.</td>
<td>No correction for age.</td>
<td>No.</td>
<td>Smears were re-examined and records of gynaecologists and gynaecological hospitals were checked. (*)</td>
<td>Yes. 10 years.</td>
<td>No statement.</td>
<td>4/9</td>
</tr>
<tr>
<td>Lyng E, 1989</td>
<td>Somewhat representative of the average Danish female population between 30 and 59 years old as some counties were excluded. (*)</td>
<td>Drawn from a different source. Different areas, with low smear-taking activity</td>
<td>No description of method. Authors only mention that data are available.</td>
<td>No.</td>
<td>The study corrected for 6 5-year age groups using a Poisson model. (*)</td>
<td>The study corrected for 4 5-year calendar periods and 19 counties. (*)</td>
<td>Record linkage with death certificates and the Danish Cancer Registry. (*)</td>
<td>Yes ~15 years.</td>
<td>No statement.</td>
<td>5/9</td>
</tr>
<tr>
<td>Magnus K, 1987</td>
<td>Truly representative of the average female population in the county of Oxford between 25 and 59 years old as all were invited for screening. (*)</td>
<td>Drawn from a different source. Female population of neighbouring counties. Rates are only used from 1963 to 1967.</td>
<td>Secure record. Screening history by national identification number. (*)</td>
<td>No.</td>
<td>No.</td>
<td>No.</td>
<td>Record linkage between the Cancer Registry and the Central Bureau of Statistics. (*)</td>
<td>Yes. 24 years.</td>
<td>No statement.</td>
<td>4/9</td>
</tr>
<tr>
<td>Mahlick C, 1994</td>
<td>Truly representative of the average Swedish female population between 30 and 49 years old as all were invited for screening. (*)</td>
<td>Drawn from the same community as the exposed cohort. Same population in the years before implementation of screening. (*)</td>
<td>No description.</td>
<td>No.</td>
<td>Yes, with the population of Sweden 1970 as a reference. (*)</td>
<td>Study controls for period and county. (*)</td>
<td>Record linkage between the Population Register (<em>) and the Cause of Death Register at the Swedish Central Bureau of Statistics. (</em>)</td>
<td>Yes. 22 years.</td>
<td>No statement.</td>
<td>6/9</td>
</tr>
</tbody>
</table>

(*) The presence of this symbol means the study fitted the selected criteria and it was accounted in the final result.
Appendix Table D
Conflict of interest and/or statements of all included studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Conflict of interest and/or funding statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bergström S, 1999</td>
<td>No statement.</td>
</tr>
<tr>
<td>Dugué P, 2014</td>
<td>Pierre-Antoine Dugué declares no potential conflict of interest. Elsebeth Lynghe and Matejka Rebolj are currently undertaking a comparative study of new-generation HPV tests, involving collaboration with Roche Diagnostics. Genomica, Qiagen and Genprobe. Elsebeth Lynghe has served as unpaid scientific advisor to Genprobe and Norchip. Matejka Rebolj’s employer received honoraria for lectures from Qiagen on her behalf. Concerning the present paper, there has been no collaboration with, or support from any of the companies. Grant sponsor: University of Copenhagen, Danish Strategic Research Council.</td>
</tr>
<tr>
<td>Ebeling K, 1986</td>
<td>No statement.</td>
</tr>
<tr>
<td>Landy R, 2016</td>
<td>The authors declare no conflict of interest. This work was supported by Cancer Research UK (A16892 to P.S.).</td>
</tr>
<tr>
<td>Lonnberg S, 2013</td>
<td>Grant sponsors: Finnish Cancer Organisations; European Union Seventh Framework Programme contract EUROOCOURSE-Europe against Cancer: Optimisation of the Use of Registries for Scientific Excellence in research. Co-funding was provided by the European Union Public Health Programme (Project no. 2006322, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines [ECCG]).</td>
</tr>
<tr>
<td>Lynge E, 1989</td>
<td>The costs of publication of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.</td>
</tr>
<tr>
<td>Magnus K, 1987</td>
<td>No statement.</td>
</tr>
<tr>
<td>Mählck C, 1994</td>
<td>This investigation was supported by Lions’ Research Foundation, University of Umeå and by Swedish Cancer Society (RmC), project no. 1759-B59-03XB.</td>
</tr>
</tbody>
</table>

Appendix E. Reference list of excluded articles


References


[29] IARC. Cancer incidence in five continents.