PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>RANDOMIZED HEALTH CARE POLICY EVALUATION OF ORGANISED PRIMARY HUMAN PAPILLOMAVIRUS (HPV) SCREENING OF WOMEN AGED 56 - 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Lamin, Helena; Eklund, Carina; Elfström, Klara; Carlsten-Thor, Agneta; Hortlund, Maria; Elfgren, Kristina; Törnberg, Sven; Dillner, Joakim</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Sveinung Sorbye</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University Hospital of North Norway, Norway</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>08-Nov-2016</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

Cervical cancer is the fourth most common cancer affecting women worldwide. Cervical cancer can be prevented by early detection and treatment of precancerous lesions. The rationale of cervical cancer screening is to identify and treat high-grade cervical intraepithelial neoplasia (CIN2+) in order to prevent its progression to invasive cancer. Since the introduction of the program, the coverage of women taking a Pap-smear has increased and consequently, the rate of cervical cancer is reduced.

Despite a well organized screening program, still women die of cervical cancer. This can be explained by a generally low clinical sensitivity of cytology and by subjectivity and poor reproducibility.

Genital human papillomavirus (HPV) is a common virus. Most sexually active people, 70-80%, will have HPV at some time in their lives. There are more than 40 types of HPV that are passed on through sexual contact. These types can infect the genital areas, including cervix, vulva, vagina, penis or anus. They can also infect the mouth and throat. Given the strong aetiological association between HPV infection and cervical cancer, HPV testing is considered an alternative for cytology-based cervical cancer screening (Ronco 2014).

When a person is exposed to HPV, the immune system usually prevents the virus from doing any serious harm. In 90% of the cases the infection is cleared within 6-24 months. But in a small number of people (10%), the virus survives for years (persistent infection). Eventually, the virus can lead to the conversion of normal cells into cancerous cells.

Even though CIN2+ is regarded as a clinically relevant lesion, and most guidelines recommend treatment of CIN2+, a considerable number of CIN2 and even CIN3 will regress without causing cervical cancer. More than 40% of CIN2 will regress within 1-2 years (Castle
2009, Guedes 2010, Ho 2011, Moscicki 2010), but CIN2 caused by HPV type 16 have a slower regression rate (Castle 2009).

A CIN1 diagnosis does not represent a significant risk factor for CIN3 above the risk attributed to its molecular cause, genotype-specific HPV infection (Castle 2011). The risk of CIN2+ in follow-up of women with negative cervical biopsy is similar for CIN0 (Normal) and CIN1 histology (Sorbye 2011). CIN1 should not be a target of screening and CIN1 should not be treated. It is not interesting to report CIN0 and CIN1 separately. In both cases the biopsy did not confirm high grade neoplasia (CIN2+).

Comments

1. A 34.5% (14763/42752) coverage of the screening program is low, especially when the women got an invitation with time and place for appointment with a midwife. I thought the coverage of the screening program in Sweden was close to 80%. Do you have any data of the non-responders? Did they skip screening at all, or do they prefer to have cytology elsewhere. Did the non-responders receive a reminder or new invitation after 12 months? What was the detection rate of CIN2+ in women not attending the study?

2. Results of screening should be reported according to the guidelines and the screening algorithm. This is the primary endpoint. Results of biopsies taken outside the screening algorithm could be reported as a secondary finding.

3. The primary evaluation was the sensitivity for CIN2+, but the absolute sensitivity cannot be estimated without biopsies from all women. It is possible to make an estimate corrected for verification bias taking random biopsies from 10% of the screening negative women in each arm. It is possible to find the relative sensitivity for CIN2+ in the HPV-arm versus the cytology-arm. The correct term is probably to evaluate the detection rate of CIN2+ in each arm.

4. You state that you report the prevalence of HPV, but the true prevalence of HPV is much higher than the detection rate (HPV-test positivity rate) using Roche Cobas 4800 because this screening test has a high cut-off to detect clinically relevant HPV-infections.

5. When reporting percentages, use one decimal and fraction, for example the detection rate of HPV 16 was 1.0% (73/7325).

6. In addition to CIN2+, the number of CIN3 and cancers in the cytology-arm and HPV-arm should be reported. One concern from the Ronco meta-analysis was the lower detection rate of cervical cancers in the first screening round.

Minor revisions

Abstract, page 2, line 12, «screening with cervical cytology (old policy)» => «screening with cervical cytology with HPV-test in triage of ASC-US / LSIL (old policy)»

Abstract, page 2, line 12-14, «screening with HPV testing, followed by triage with cytology (new policy)» => screening with HPV-test with cytology in triage of HPV positive (new policy)

Abstract, page 2, line 24-26, «In the new policy, the population HPV
prevalence was 5.5%» => «In the new policy, the detection rate of HPV using Roche Cobas 4800 was 5.5% (405/7325)»

Abstract, page 2, line 26-28, delete «There was thus a 94.5% decline in number of cytologies performed»

Abstract, page 2, line 28, «HPV16 prevalence was 1.0% and HPV18 prevalence was 0.3%» => «The detection rates of HPV16 and 18 were 1.0% (73/7325) and 0.3% (22/7325), respectively»

Abstract, page 2, line 28-30, «In the HPV policy arm, 78/405 (19%) HPV-positive women were also cytology positive and there were 19 cases of CIN2+ in histopathology» => «In the HPV policy arm, 19.3% (78/405) were triage positive, 73.1% (57/78) had biopsy and there were 19 cases of CIN2+, PPV 33.3% (19/57)»

Abstract, page 2, line 28-31, «In the cytology policy, 153 women were cytology positive and there were 18 cases of CIN2+ in histopathology» => «In the cytology policy, 2.1% (153/7438) were cytology positive, 40.5% (62/153) were triage positive, 82.3% (51/62) had biopsy and there were 16 cases of CIN2+, PPV 31.4% (16/51)»

Abstract, page 2, line 32-34, add «The referral rates for colposcopy in the HPV-arm and the cytology-arm were 1.1% (78/7325) and 0.8% (62/7438), respectively. The detection rates of CIN2+ were 0.3% (19/7325) and 0.2% (16/7438)»

Abstract, page 2, line 34-39, delete «Both the totalt number of cervical biopsies and the number of cervical biopsies with benign histopathology was much lower in the HPV policy (49 benign, 87 total versus 105 benign, 132 total»

Abstract, page 2, line 41-45, «primary HPV screening had, already before follow-up of HPV-positive women with new HPV test and referral of women with persistence, a similar sensitivity for CIN2+ as cytology-based screening» => «primary HPV screening had, already before follow-up of HPV-positive, triage negative women, with new HPV test, and referral of women with persistence, a similar detection rate for CIN2+ as cytology-based screening»

Introduction, page 4, line 30-32, «An HPV test before exiting the program is expected to give a longer lasting protection for cervical cancer than a cytology test» => «A negative HPV test before exiting the program is expected to give a longer lasting protection for cervical cancer than a negative cytology test»

Material and methods, cytology policy, page 7, line 5, «Women with CIN2+ in primary cytology were always referred to colposcopy» => «Women with high grade cytology (ASC-H / HSIL / ACIS) were referred to colposcopy»

Material and methods, cytology policy, page 7, line 5-7, «Those with low grade cytology (ASC-US/LSIL) were triaged with HPV-testing»

Results, page 8, line 11-26, see my suggested corrections in the abstract

Discussion, page 9, line 23-25, «The exact population-based HPV
prevalences in this age group (with the HPV test used) was not known before, but the 5.5% prevalence was approximately as expected.» Add «In the Horizon study, across all age groups, 898 of 4413 (20.3%) women with normal cytology were HPV DNA-positive (Roche Cobas 4800). This proportion decreased from 33.9% in women aged 25–29 years to 18.6% at age 30–39 years and 6.0% at age 50–65 years (Preisler 2013).»

Discussion, page 9, line 27-29, «The HPV test is performed using laboratory automation and the amount of personnel time required and costs was substantially reduced.» I am not sure if the price of a HPV-test is low enough to be compared with a screener. How many cytology samples can be screened manually during one day? In Norway a screener usually screen 30-35 samples a day. If the price of a HPV-test is SEK 100, the wages to a screener have to be higher than SEK 3,000-3,500 a day to be less cost-effective than HPV-testing with the same screening interval. The real cost-saving is extending the screening intervals from cytology every three year to HPV-testing every five years, or from cytology every five years to HPV-testing every ten years.

Discussion, page 9, line 32-34, «The number of primary referrals for gynecological follow-up was substantially reduced (to about half)». This is only true if you also include the cytology samples taken outside the cervical screening program. Endometrioid adenocarcinoma of the uterus and dysplasia in the vagina or vulva is not a target for the screening program.

Discussion, page 11, add the same conclusion as in the abstract

Table 1, page 13, after «Attending women» add «screening positive women» and «triage positive women»

Table 1, page 13, «Referral rate to histopathology» => «Women with biopsy»

Table 1, page 13, women with CIN2+ in cytology arm «18» => «16»

Table 1, page 13, delete «Lowgrade lesions» and «Benign histopathology»

Table 1, page 13, add «Women with CIN3» and «Women with cervical cancer»

Table 2, page 14, delete «Lowgrade lesions (HPV positive / All)» and «Benign histopathology (HPV positive / All)»

Table 3, page 14, delete «Lowgrade lesions (Abnormal cytology / All)» and «Benign histopathology (Abnormal cytology / All)»

Figure 1, page 15, add the box «Triage positive: 62» between «Abnormal cytology: 153» and «PAD: 51» in the cytology arm

References


Castle PE, Schiffman M, Wheeler CM, Solomon D. Evidence for
This interesting randomized study compares the sensitivity for CIN2+ of an organized primary HPV screening program (primary HPV policy) with the current cytology screening program (primary cytology policy) in women 56-60 year in a real life screening setting in the Stockholm-Gotland region in Sweden taking into account all aspects of the currently running cytology screening program. HPV pos women are triaged by cytology and if cytology was ≥ ASC-US-CIN1/LSIL women were referred for colposcopy.

In the HPV policy arm HPV pos women with normal (NILM) cytology were intended to be followed by HPV and cytology initially after 1 year which was later extended to 3 years. Therefore the results of the f-up of these HPV pos women with normal cytology are not part of the CIN2+ evaluation of the present paper and make the comparison the status an interim report.

The results of this randomized study in old 56-60y) women show that already without follow-up of HPV pos cytology negative women the sensitivity for CIN2+ in the HPV arm is equal to that of cytology. Moreover the total number of cervical biopsies and the number of...
cervical biopsies with benign histopathology was much lower in the HPV policy arm than in the cytology arm.

Remarks

1. An additional surprising finding is that the number of histopathological biopsies in the cytology arm (n=132) was higher than in the HPV arm (n=85). In the last paragraph on page 10 the authors try to explain that the relatively high number of biopsies detected in cytology negative women (n=85) of the cytology policy arm is due to screening for vulva, vagina and endometrium and brings up the question whether this is common practice, also in cervical cytology screening programs in other countries. If the answer of this question is that it is an incidental non-structural finding then this may influence the remark in the discussion by the authors that the number of referrals in the HPV pos arm is lower.(page9 line13-14)

2. The CIN2+ should be further specified into CIN2, CIN3 and Cancer. Also the catagory "benign histopathologies should be further specified.

3. Page 10 line10-11 I do not understand the sentence "the fact that twice as many smears with ASC-US/LSIL diagnosis were detected in the HPV arm screening arm may be a reflection of this" In the cytology arm are 134 ASC-US/LSIL cases but the figure for the HPV arm is not given. Please explain

4. "Benign histopathologies (HPV pos )" should be specified. According to table 2 they are not CIN1 lesions? What is the difference with benign histopathologies (abnormal cytology) in table 3 in the cytology arm. Same for low grade lesions (HPV positive) in the HPV arm and low grade lesions in the abnormal cytology arm. Are the same f-up policies used for the HPV pos low grade and benign histopathologies as for the ones detected by cytology"? This may be important for future costs calculations

5. Check in flow chart (fig1) HPV abnormal 78 but there are only 57 biopsies taken. What happened with the 21 women? . Same story in the cytology arm : 153 cytology abnormal ,of which 134 ASCUS/LSIL, of which 43 HPV pos and thus got a biopsy (n=43) . Women with abnormal cytology 153-134 -ASCUS/LSIL= 19 women should have high SIL and thus should have a biopsy. 43+19= 62 Women should have a biopsy: what happened to the 11 women who should have had a biopsy.

5. A few more references might be given in the discussion.
equivocal cytology results (ASC-US in the Bethesda System) are found?

2. The sensitivity of cervical screening decreases post-menopausally due to increased inadequacy of sampling and difficulty diagnosing women with small lesions beyond visualization upon referral (i.e., inadequate colposcopy). Please discuss what brush was used and its ability to collect cells at the SCJ in this age group. When tests were positive, how many colposcopic examinations were inadequate due to failure to see the SCJ, and what was done when inadequate? Do the authors believe that CIN2+ cases were missed as a result?

3. This is a randomized trial of primary HPV screening with cytology triage compared with the current norm of cytology screening, among women 56-60. The same specimen is collected and testing by both tests could be used on all women to increase power if funding permits. I presume the reason for randomization is to create a population effectiveness trial, and that assessment of efficacy of the 2 test methods was not the issue given many previous comparisons. The rationale could be made more clear in the text. The small numbers of CIN2+ led to a descriptive analysis without significance testing. A paired analysis with testing of all specimens would have produced a more powerful comparison, this could be noted.

4. The participation rate of around 35% was quite low, below the level at which we can comfortably assume that the findings are generalizable to all women in the 56-60 year old population. Please address this issue. In terms of population effectiveness, both arms could be judged to be inadequate for final assessment prior to exiting the screening program.

5. It is my belief that we lack adequate data to know the performance of HPV testing for exiting women from screening, because the needed data will require prospective follow-up of cancer occurrence post exiting in the HPV era. The current U.S. guidelines for exiting (cytology or HPV or cotest-based) are not sufficiently evidence-based, in my view. What is the opinion of the authors regarding the negative predictive value of 1 or 2 negative primary HPV tests, with previous history of negative screening, with regard to subsequent cancer risk? What if the woman does not have a prior history of a lifetime of negative screening? Is her probability of cancer low enough given 1 or 2 negative HPV tests to justify never screening her again?

6. The authors are correct in stating that a limitation of this analysis is that it reports a truncated screening round. HPV positive cytology negative women and HPV negative cytology ASC-US women are seen back at 1-3 years, and they represent (especially the HPV positive cytology negative ones) a non-negligible fraction of CIN2+ to be found at that complete round. Please comment what the expectations of complete screening might be, based on data in the literature. Also, repeat positivity of cobas results is an inexact measure of HPV persistence. Is an HPV16 followed by HR12 positive considered double-positive, given the obvious type difference? And what about HR12 twice, what percentage are truly HPV persistence, please discuss within the limitations of known data.

7. Overall, my point is that we are entering a new era of
screening with inadequate knowledge of how HPV testing will perform especially in exiting at 60 or 65 based on 1 or 2 negative tests. That does not mean HPV testing should not be adopted, just that exiting has always been tricky and it remains to be seen how HPV testing affects its safety. I believe this limitation is not sufficiently recognized by guidelines groups. What do the authors say? It seems that assurance of adequate screening and examination of the SCJ might increase the reassurance of a negative screen prior to exiting, as the other general concern.

REVIEWER
David Collett
NHS Blood and Transplant
UK

REVIEW RETURNED
12-Dec-2016

GENERAL COMMENTS
This study is based on women from a narrow age range (56-60) and is highly self-selected. Thus of 42752 women approached, only 14763 participated. This may well be a group that is different in many ways from the general population, especially in educational attainment and social status. For this reason, the results cannot be used to infer about a wider population. The authors note that the results do not yet enable any conclusions to be drawn about the longer term benefits of HPV screening before reaching 60, another aim of the study. There is also no mention of a power calculation, so we cannot know whether the study would have had sufficient power to distinguish between the results from the two groups.

Accepting that the data can only be used to compare the two procedures in a highly selected sample, the stated aim is to compare the sensitivities of the two approaches. Sensitivity is of course the proportion of true positives correctly identified, and so to determine this, we need to know the CIN2+ status of the 14763 participants. However, we only know this status for the women who had PAD, of which there were 19/85 in the Primary HPV group and 18/132 in the Primary cytology group. Comparing these proportions does not help in the assessment of the sensitivities of the overall procedures. It is reassuring that the two procedures identify about the same proportion of CIN2+ women randomised (19/7325 and 18/7438), but these figures do not relate to a general population.

Accepting that there is relevant information in the summary statistics given, it is nevertheless not easy to reconcile the data in Tables 1 – 3 with the results and discussion sections. Figure 1 certainly helps, but I think the denominators of the summary proportions need to be more clearly defined so that the value of the results in the tables can be evaluated. The results section can then be better aligned with the tables and the stated aims of the study.

There is no statistical procedures are used to make any comparisons between the results from the two groups and there are no interval estimates for the proportions given.

A number of acronyms have not been spelt out at first occurrence, including ASCUS, LSIL and RHP.
Cervical cancer is the fourth most common cancer affecting women worldwide. Cervical cancer can be prevented by early detection and treatment of precancerous lesions. The rationale of cervical cancer screening is to identify and treat high-grade cervical intraepithelial neoplasia (CIN2+) in order to prevent its progression to invasive cancer. Since the introduction of the program, the coverage of women taking a Pap-smear has increased and consequently, the rate of cervical cancer is reduced.

Despite a well organized screening program, still women die of cervical cancer. This can be explained by a generally low clinical sensitivity of cytology and by subjectivity and poor reproducibility.

Genital human papillomavirus (HPV) is a common virus. Most sexually active people, 70-80%, will have HPV at some time in their lives. There are more than 40 types of HPV that are passed on through sexual contact. These types can infect the genital areas, including cervix, vulva, vagina, penis or anus. They can also infect the mouth and throat. Given the strong aetiological association between HPV infection and cervical cancer, HPV testing is considered an alternative for cytology-based cervical cancer screening (Ronco 2014).

When a person is exposed to HPV, the immune system usually prevents the virus from doing any serious harm. In 90% of the cases the infection is cleared within 6-24 months. But in a small number of people (10%), the virus survives for years (persistent infection). Eventually, the virus can lead to the conversion of normal cells into cancerous cells.

Even though CIN2+ is regarded as a clinically relevant lesion, and most guidelines recommend treatment of CIN2+, a considerable number of CIN2 and even CIN3 will regress without causing cervical cancer. More than 40% of CIN2 will regress within 1-2 years (Castle 2009, Guedes 2010, Ho 2011, Moscicki 2010), but CIN2 caused by HPV type 16 have a slower regression rate (Castle 2009). A CIN1 diagnosis does not represent a significant risk factor for CIN3 above the risk attributed to its molecular cause, genotype-specific HPV infection (Castle 2011). The risk of CIN2+ in follow-up of women with negative cervical biopsy is similar for CIN0 (Normal) and CIN1 histology (Sorbye 2011). CIN1 should not be a target of screening and CIN1 should not be treated. It is not interesting to report CIN0 and CIN1 separately. In both cases the biopsy did not confirm high grade neoplasia (CIN2+).

We agree on this summary of the field, which emphasizes the importance of the subject.

Comments

1. A 34.5% (14763/42752) coverage of the screening program is low, especially when the women got an invitation with time and place for appointment with a midwife. I thought the coverage of the screening program in Sweeden was close to 80%. Do you have any data of the non-responders? Did they skip screening at all, or do they prefer to have cytology elsewhere. Did the non-responders receive a reminder or new invitation after 12 months? What was the detection rate of CIN2+ in women not attending the study?

2. Results of screening should be reported according to the guidelines and the screening algorithm. This is the primary endpoint. Results of biopsies taken outside the screening algorithm could be reported as a secondary finding.
The study was performed entirely within the real-life screening program. This is now better explained.

3. The primary evaluation was the sensitivity for CIN2+, but the absolute sensitivity cannot be estimated without biopsies from all women. It is possible to make an estimate corrected for verification bias taking random biopsies from 10% of the screening negative women in each arm. It is possible to find the relative sensitivity for CIN2+ in the HPV-arm versus the cytology-arm. The correct term is probably to evaluate the detection rate of CIN2+ in each arm.

*We have introduced the terminology recommended by the reviewer (Detection rate of CIN2+).

4. You state that you report the prevalence of HPV, but the true prevalence of HPV is much higher than the detection rate (HPV-test positivity rate) using Roche Cobas 4800 because this screening test has a high cut-off to detect clinically relevant HPV-infections.

*The issue of whether infections that can not be detected still exist is hard to address experimentally. We have rephrased the statement to read “prevalence of HPV, using an accredited HPV test (Cobas 4800)”.

5. When reporting percentages, use one decimal and fraction, for example the detection rate of HPV 16 was 1.0% (73/7325).

*Changed as requested.

6. In addition to CIN2+, the number of CIN3 and cancers in the cytology-arm and HPV-arm should be reported. One concern from the Ronco meta-analysis was the lower detection rate of cervical cancers in the first screening round.

*This is now detailed, in Figure 1.

Minor revisions
Abstract, page 2, line 12, «screening with cervical cytology (old policy)» => «screening with cervical cytology with HPV-test in triage of ASC-US / LSIL (old policy)»
*Changed as requested.

Abstract, page 2, line 12-14, «screening with HPV testing, followed by triage with cytology (new policy)» => screening with HPV-test with cytology in triage of HPV positive (new policy)»
*Changed as requested.

Abstract, page 2, line 24-26, «In the new policy, the population HPV prevalence was 5.5%» => «In the new policy, the detection rate of HPV using Roche Cobas 4800 was 5.5% (405/7325)»
*Changed to emphasize the HPV test used.

Abstract, page 2, line 26-28, delete «There was thus a 94.5% decline in number of cytologies performed»
*Changed as requested.

Abstract, page 2, line 28, «HPV16 prevalence was 1.0% and HPV18 prevalence was 0.3%» => «The detection rates of HPV16 and 18 were 1.0% (73/7325) and 0.3% (22/7325), respectively»
*Changed as requested.

Abstract, page 2, line 28-30, «In the HPV policy arm, 78/405 (19%) HPV-positive women were also cytology positive and there were 19 cases of CIN2+ in histopathology» => «In the HPV policy arm, 19.3% (78/405) were triage positive, 73.1% (57/78) had biopsy and there were 19 cases of CIN2+,
PPV 33.3% (19/57)

*Rephrased to make more clear.

Abstract, page 2, line 32-34, «In the cytology policy, 153 women were cytology positive and there were 18 cases of CIN2+ in histopathology» => «In the cytology policy, 2.1% (153/7438) were cytology positive, 40.5% (62/153) were triage positive, 82.3% (51/62) had biopsy and there were 16 cases of CIN2+, PPV 31.4% (16/51)»

Abstract, page 2, line 34, add «The referral rates for colposcopy in the HPV-arm and the cytology-arm were 1.1% (78/7325) and 0.8% (62/7438), respectively. The detection rates of CIN2+ were 0.3% (19/7325) and 0.2% (16/7438)»

Abstract, page 2, line 34-39, delete «Both the totalt number of cervical biopsies and the number of cervical biopsies with benign histopathology was much lower in the HPV policy (49 benign, 87 total versus 105 benign, 132 total»

The reviewer proposes changing the actual number of counts in the study and deletion of some of the data. We are of course eager to improve the manuscript, but we can not change actual results.

Abstract, page 2, line 41-45, «primary HPV screening had, already before follow-up of HPV-positive women with new HPV test and referral of women with persistence, a similar sensitivity for CIN2+ as cytology-based screening» => «primary HPV screening had, already before follow-up of HPV-positive, triage negative women, with new HPV test, and referral of women with persistence, a similar detection rate for CIN2+ as cytology-based screening»

Changed as requested.

Introduction, page 4, line 30-32, «An HPV test before exiting the program is expected to give a longer lasting protection for cervical cancer than a cytology test» => «A negative HPV test before exiting the program is expected to give a longer lasting protection for cervical cancer than a negative cytology test»

*Changed as requested.

Material and methods, cytology policy, page 7, line 5, «Women with CIN2+ in primary cytology were always referred to colposcopy» => «Women with high grade cytology (ASC-H / HSIL / ACIS) were referred to colposcopy»

*The terminology used was CIN2+. This is indeed approximately the same as ASC-H / HSIL / ACIS, but as we did not use the ASC-H / HSIL / ACIS terminology in the study, we can not introduce this terminology here.

Material and methods, cytology policy, page 7, line 5-7, «Those with low grade cytology (ASC-US/LSIL) were triaged with HPV-testing»

*Changed as requested.

Results, page 8, line 11-26, see my suggested corrections in the abstract

Rephrasing to make the data more clear.

Discussion, page 9, line 23-25, «The exact population-based HPV prevalences in this age group (with the HPV test used) was not known before, but the 5.5% prevalence was approximately as expected.» Add «In the Horizon study, across all age groups, 898 of 4413 (20.3%) women with normal cytology were HPV DNA-positive (Roche Cobas 4800). This proportion decreased from 33.9% in women aged 23–29 years to 18.6% at age 30–39 years and 6.0% at age 50–65 years (Preisler 2013).»

*Changed as requested.

Discussion, page 9, line 27-29, «The HPV test is performed using laboratory automation and the amount of personnel time required and costs was substantially reduced.» I am not sure if the price of
a HPV-test is low enough to be compared with a screener. How many cytology samples can be screened manually during one day? In Norway a screener usually screen 30-35 samples a day. If the price of a HPV-test is SEK 100, the wages to a screener have to be higher than SEK 3,000-3,500 a day to be less cost-effective than HPV-testing with the same screening interval. The real cost-saving is extending the screening intervals from cytology every three year to HPV-testing every five years, or from cytology every five years to HPV-testing every ten years.

*The meaning of this comment is not entirely clear to us. It states that a cytology screener screens 30-35 samples per day, arguing that this method must be less expensive. By comparison, an HPV technician analyses about 270 samples per day. We did not itemize any expenses as this is not a health economics study, we merely noted that the total expenses went down. We have inserted a word of caution in interpretation.

Discussion, page 9, line 32-34, «The number of primary referrals for gynecological follow-up was substantially reduced (to about half)». This is only true if you also include the cytology samples taken outside the cervical screening program. Endometrioid adenocarcinoma of the uterus and dysplasia in the vagina or vulva is not a target for the screening program.

*We included only samples from the cervix, not from the uterus, vagina or vulva. This is now emphasized.

Discussion, page 11, add the same conclusion as in the abstract.

*Changed as requested.

Table 1, page 13, after «Attending women» add «screening positive women» and «triage positive women»

Table 1, page 13, «Referral rate to histopathology» => «Women with biopsy»

Table 1, page 13, women with CIN2+ in cytology arm «18» => «16»

Table 1, page 13, delete «Lowgrade lesions» and «Benign histopathology»

Table 1, page 13, add «Women with CIN3» and «Women with cervical cancer»

Table 2, page 14, delete «Lowgrade lesions (HPV positive / All)» and «Benign histopathology (HPV positive / All)»

Table 3, page 14, delete «Lowgrade lesions (Abnormal cytology / All)» and «Benign histopathology (Abnormal cytology / All)»

Figure 1, page 15, add the box «Triage positive: 62» between «Abnormal cytology: 153» and «PAD: 51» in the cytology arm

*As discussed above, we find it problematic to change the actual counts of the study, as presented in the Tables.

References


Reviewer: 2
Reviewer Name: Chris JLM.Meijer
Institution and Country: Dept of Pathology, Vrije Universiteit Medical center, Amsterdam, Te Netherlands
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This interesting randomised study compares the sensitivity for CIN2+ of an organized primary HPV screening program (primary HPV policy) with the current cytology screening program (primary cytology policy) in women 56-60 year in a real life screening setting in the Stockholm-Gotland region in Sweden taking into account all aspects of the currently running cytology screening program. HPV pos women are triaged by cytology and if cytology was ≥ ASC-US-CIN1/LSIL women were referred for colposcopy.

In the HPV policy arm HPV pos women with normal (NILM) cytology were intended to be followed by HPV and cytology initially after 1 year which was later extended to 3 years. Therefore the results of the f-up of these HPV pos women with normal cytology are not part of the CIN2+ evaluation of the present paper and make the comparison the status an interim report.

The results of this randomized study in old 56-60y) women show that already without follow-up of HPV pos cytology negative women the sensitivity for CIN2+ in the HPV arm is equal to that of cytology. Moreover the total number of cervical biopsies and the number of cervical biopsies with benign histopathology was much lower in the HPV policy arm than in the cytology arm.

*We thank the reviewer for this supportive comment.

Remarks

1. An additional surprising finding is that the number of histopathological biopsies in the cytology arm (n=132) was higher than in the HPV arm (n=85). In the last paragraph on page 10 the authors try to explain that the relatively high number of biopsies detected in cytology negative women (n=85) of the cytology policy arm is due to screening for vulva, vagina and endometrium and brings up the question whether this is common practice, also in cervical cytology screening programs in other countries. If the answer of this question is that it is an incidental non-structural finding then this may influence the remark in the discussion by the authors that the number of referrals in the HPV pos arm is lower.(page9 line13-14)

*We have inserted a word of caution that we do not know the generalizability of this specific finding and that this is a research priority for the future.

2. The CIN2+ should be further specified into CIN2, CIN3 and Cancer. Also the catagory "benign histopathologies should be further specified.
This was also requested by reviewer 1 and has been added to Figure 1. CIN2+ was the prespecified hypothesis of the study and sub-splitting it results in low numbers. However, we decided to present the exact sub-splitted counts as people doing systematic reviews may need also small exact counts. We now also further explain the meaning of the category “benign histopathologies”.

3. Page 10 line10-11 I do not understand the sentence “the fact that twice as many smears with ASC-US/LSIL diagnosis were detected in the HPV arm screening arm may be a reflection of this” In the cytology arm are 134 ASC-US/LSIL cases but the figure for the HPV arm is not given. Please explain.

“We have rephrased to make this more clear.

4.”Benign histopathologies (HPV pos)” should be specified. According to table 2 they are not CIN1 lesions? What is the difference with benign histopathologies (abnormal cytology) in table 3 in the cytology arm. Same for low grade lesions (HPV positive) in the HPV arm and low grade lesions in the abnormal cytology arm. Are the same f-up policies used for the HPV pos low grade and benign histopathologies as for the ones detected by cytology”? This may be important for future costs calculations

“As stated above, we have now explained the term benign histopathologies. We have also inserted a sentence to emphasize that the follow-up policies used were exactly the same.

4. Check in flow chart (fig1) HPV abnormal 78 but there are only 57 biopsies taken. What happened with the 21 women? .

Same story in the cytology arm : 153 cytology abnormal ,of which 134 ASCUS/LSIL, of which 43 HPV pos and thus got a biopsy (n=43) . Women with abnormal cytology 153-134 -ASCUS/LISL= 19 women should have high SIL and thus should have a biopsy. 43+19= 62 Women should have a biopsy: what happened to the 11 women who should have had a biopsy.

“We have inserted a paragraph in the discussion explaining the difference between a real-life study and a research study.

5. A few more references might be given in the discussion.

T"his has been added.

Reviewer: 3
Reviewer Name: Mark Schiffman

1. Please explain the current recommendations for exiting (cytology) screening invitations in this area of Sweden. What longitudinal combinations lead to exiting, and how are other women followed? For example, are 2 sequential negative results sufficient for exiting, and what happens if equivocal cytology results (ASC-US in the Bethesda System) are found?

“We have inserted a paragraph to explain this (the rules are same for the last screening test as for all other screening tests).

2. The sensitivity of cervical screening decreases post-menopausally due to increased inadequacy of sampling and difficulty diagnosing women with small lesions beyond visualization upon referral (i.e., inadequate colposcopy). Please discuss what brush was used and its ability to collect cells at the SCJ in this age group. When tests were positive, how many colposcopic examinations were inadequate due to failure to see the SCJ, and what was done when inadequate? Do the authors believe that CIN2+ cases were missed as a result?

“We have added description of the brush used, we cite previous work on the management algorithm used in case of inadequate colposcopy and we reference previous work of the very high cancer-protective effect of the program (which argues that the policy used is basically adequate).

3. This is a randomized trial of primary HPV screening with cytology triage compared with the current norm of cytology screening, among women 56-60. The same specimen is collected and testing by both tests could be used on all women to increase power if funding permits. I presume the reason for
randomization is to create a population effectiveness trial, and that assessment of efficacy of the 2 test methods was not the issue given many previous comparisons. The rationale could be made more clear in the text. The small numbers of CIN2+ led to a descriptive analysis without significance testing. A paired analysis with testing of all specimens would have produced a more powerful comparison, this could be noted.

*We have added a paragraph to elaborate between the difference between studies in the research setting and population effectiveness trials. In short, we think there are so many research studies that another one (e.g. using double testing) would not be likely to detect anything novel. By contrast, there are very few real-life effectiveness studies.

4. The participation rate of around 35% was quite low, below the level at which we can comfortably assume that the findings are generalizable to all women in the 56-60 year old population. Please address this issue. In terms of population effectiveness, both arms could be judged to be inadequate for final assessment prior to exiting the screening program.

*This issue was also raised by reviewer 4 and is a misunderstanding. We have tried to explain more clearly that the population coverage was 74.4%, which is very high by international comparison. The metric “attendance rate” measures attendance after each invitation and is much lower, because of systematic issuing of reminder invitations to women who did not attend the appointment in their previous invitation.

5. It is my belief that we lack adequate data to know the performance of HPV testing for exiting women from screening, because the needed data will require prospective follow-up of cancer occurrence post exiting in the HPV era. The current U.S. guidelines for exiting (cytology or HPV or cotest-based) are not sufficiently evidence-based, in my view. What is the opinion of the authors regarding the negative predictive value of 1 or 2 negative primary HPV tests, with previous history of negative screening, with regard to subsequent cancer risk? What if the woman does not have a prior history of a lifetime of negative screening? Is her probability of cancer low enough given 1 or 2 negative HPV tests to justify never screening her again?

*We agree that this is an important and urgent issue to study. The issue requires careful and extensive analysis of comprehensive registries and will be the subject of another paper.

6. The authors are correct in stating that a limitation of this analysis is that it reports a truncated screening round. HPV positive cytology negative women and HPV negative cytology ASC-US women are seen back at 1-3 years, and they represent (especially the HPV positive cytology negative ones) a nonnegligible fraction of CIN2+ to be found at that complete round. Please comment what the expectations of complete screening might be, based on data in the literature. Also, repeat positivity of cobas results is an inexact measure of HPV persistence. Is an HPV16 followed by HR12 positive considered double-positive, given the obvious type difference? And what about HR12 twice, what percentage are truly HPV persistence, please discuss within the limitations of known data.

*We have inserted the desired comment. We have also explained that the algorithm required type-specific persistence (not only persistent Cobas positivity).

7. Overall, my point is that we are entering a new era of screening with inadequate knowledge of how HPV testing will perform especially in exiting at 60 or 65 based on 1 or 2 negative tests. That does not mean HPV testing should not be adopted, just that exiting has always been tricky and it remains to be seen how HPV testing affects its safety. I believe this limitation is not sufficiently recognized by guidelines groups. What do the authors say? It seems that assurance of adequate screening and examination of the SCJ
might increase the reassurance of a negative screen prior to exiting, as the
other general concern.
*This is the same comment as for point 5. We agree completely, but will be addressing this in another paper.
Reviewer: 4
Reviewer Name: David Collett
Please leave your comments for the authors below
This study is based on women from a narrow age range (56-60) and is highly self-selected. Thus of 42752 women approached, only 14763 participated. This may well be a group that is different in many ways from the general population, especially in educational attainment and social status. For this reason, the results cannot be used to infer about a wider population. The authors note that the results do not yet enable any conclusions to be drawn about the longer term benefits of HPV screening before reaching 60, another aim of the study. There is also no mention of a power calculation, so we cannot know whether the study would have had sufficient power to distinguish between the results from the two groups.

*We have tried to explain further that our design is extremely population-based and exactly the same as used in the organized screening program.
We have also inserted the power calculation.

Accepting that the data can only be used to compare the two procedures in a highly selected sample, the stated aim is to compare the sensitivities of the two approaches. Sensitivity is of course the proportion of true positives correctly identified, and so to determine this, we need to know the CIN2+ status of the 14763 participants. However, we only know this status for the women who had PAD, of which there were 19/85 in the Primary HPV group and 18/132 in the Primary cytology group.
Comparing these proportions does not help in the assessment of the sensitivities of the overall procedures. It is reassuring that the two procedures identify about the same proportion of CIN2+ women randomised (19/7325 and 18/7438), but these figures do not relate to a general population. Accepting that there is relevant information in the summary statistics given, it is nevertheless not easy to reconcile the data in Tables 1 – 3 with the results and discussion sections. Figure 1 certainly helps, but I think the denominators of the summary proportions need to be more clearly defined so that the value of the results in the tables can be evaluated. The results section can then be better aligned with the tables and the stated aims of the study.

*We have revised the results section to make it more clear and well aligned with the Tables.

There is no statistical procedures are used to make any comparisons between the results from the two groups and there are no interval estimates for the proportions given.

*The statistical analysis is now given, including the confidence interval for the proportions.

A number of acronyms have not been spelt out at first occurrence, including ASCUS, LSIL and RHP.

*The acronyms are now spelt out.
<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sveinung Sorbye</td>
<td>University Hospital of North Norway, Norway</td>
</tr>
<tr>
<td>Chris JLM Meijer</td>
<td>Vrije Universiteit Medical Center. dept of Pathology Amsterdam The Netherlands</td>
</tr>
<tr>
<td>Mark Schiffman</td>
<td>NCI, USA</td>
</tr>
<tr>
<td>David Collett</td>
<td>NHS Blood and Transplant, UK</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

- All comments have been addressed. The response to the comments is adequate and improved the manuscript.
- The authors have responded on my comments adequately and modified the manuscript accordingly. I have no further comments. In general the revised version of the manuscript has improved.
- This is a well-done and timely effort. In our own evaluations of HPV testing and cytology at older ages, we note some concerns with decreased sensitivity of both methods, possibly due to specimen sampling issues. In your program, you use European recommendations to extend return times for women to 3 years, despite HPV positive findings. Are the general recommendations adequate for older women? How do you address technical sensitivity concerns when the T zone is potentially missed by usual sampling?
- Thank you for responding to my earlier comments.
VERSION 2 – AUTHOR RESPONSE

Reviewers 1, 2 and 4: Had no specific comments, apart from recommending publication.
Reviewer 3: We have inserted a sentence to describe the reasoning behind the return time of 3 years, also in this age group where sampling of the T zone may be difficult.
Randomised healthcare policy evaluation of organised primary human papillomavirus screening of women aged 56–60

Helena Lamin, Carina Eklund, Klara Miriam Elfström, Agneta Carlsten-Thor, Maria Hortlund, Kristina Elfgren, Sven Törnberg and Joakim Dillner

BMJ Open 2017 7:
doi: 10.1136/bmjopen-2016-014788

Updated information and services can be found at:
http://bmjopen.bmj.com/content/7/5/e014788

These include:

References
This article cites 15 articles, 4 of which you can access for free at:
http://bmjopen.bmj.com/content/7/5/e014788#BIBL

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Public health (2287)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/