UK National Screening Committee
Consultation on modifying the NHS Cervical Screening Programmes in the four UK nations

Purpose

The UK National Screening Committee (UK NSC) recommended the use of primary HPV screening in the cervical screening programme in November 2015. The purpose of this coversheet is to set out a series of issues to help operationalise HPV based screening in the four UK nations. The UK NSC would like to hear from stakeholders on the recommendations made on each issue.

Issue 1: screening intervals and surveillance intervals

Proposed recommendation:

It is proposed that the Cervical Screening Programmes in the four UK nations should implement the following:

- an expanded, five year, screening interval for HPV negative women
- a 12 month surveillance interval for HPV positive / cytology negative women and that
- women with persistent HPV infection and negative cytology should undergo two surveillance tests. If HPV positive at the second test they should be referred to colposcopy irrespective of cytology result

Justification

The evidence base for expanded screening intervals was discussed in the previous UK NSC review. This focused on several European studies of HPV / cytology co-testing (document 2). Reports from the HPV pilot sites lent further weight to the viability of extended screening intervals. Early data from the pilot sites is also discussed in document 1.

However there is little, direct, relevant primary research evidence to guide this discussion. Because of this, modelling studies have been undertaken in the UK and internationally to explore the likely impact of extended screening intervals. In the UK, the NHS National Cervical Screening Programme (NCSP) commissioned a team within PHE to produce a model exploring the screening and surveillance intervals. This model (document 3a) estimated that a strategy as proposed by the UK NSC would result in:

- a decrease in the annual number of primary screening tests and no change in the number colposcopies
- an increase in detection of CIN2+ and a reduction of cancer incidence and cancer related deaths
- an annual reduction in health related costs and an uncertain impact on quality adjusted life years

A review of modelling studies undertaken in this area suggests that these outcomes are consistent with estimates developed as outputs from other modelling exercises (document 4).

Comments

The UK NSC would welcome comments on the proposed strategy.

Issue 2: women aged 64 and over who are exiting the programme

There is an absence of evidence to guide recommendations on women exiting the programme. For example, no estimates of outcomes in this age group were identified in the summary of modelling studies (document 4).

Proposed recommendation:

It is proposed that the Cervical Screening Programmes in the four UK nations should implement the following steps :

- HPV positive / cytology positive women should be managed in the same way as other age groups
- HPV positive / cytology negative women should be recalled at 12 months and, if still HPV positive, be referred for colposcopy. If colposcopy is:

    i. decisively negative this would prompt discharge from the programme
    ii. decisively positive this would prompt the offer of loop excision
    iii. indecisive this would prompt the offer of loop excision or recall a further 12 months later
- as there is an absence of evidence in this area the Programme should work with the relevant national professional or standard setting bodies to produce a clinical consensus statement to guide practice in this area.

Comments

The UK NSC would welcome comments on the proposed strategy.

Issue 3: Self sampling as a strategy to address non attendance for screening

Proposed recommendation

It is proposed that self sampling as a strategy to address non attendance for screening requires further study in well organised pilots and research projects.

Other questions relating to the fit between this approach and the screening programme should also be the subject of research and piloting. For example this would apply to the use of self sampling as an approach to routine screening programme delivery.

Justification

A rapid review of the evidence relating to self sampling is attached (document 5). The current draft of the document was completed in in March 2017 and reported that:

i) test performance is reasonable and may be useful as a failsafe for women who do not respond to screening invitations

ii) there is a low rate of inadequate samples for HPV testing

iii) there was an improvement in screening uptake, in most studies, of between ~10% - ~20% when compared to invitations to clinician sampling

iv) a proportion of women did not use the kits but were prompted to attend clinician sampling. This proportion varied considerably between studies

However, the review highlighted a number of limitations:

- cost effectiveness of the strategy had not been evaluated
- there was insufficient information on the circumstances in which the approach should be used. This might include the overall level of uptake, length of time following the initial invitation and the number of subsequent prompts
- the review suggested that it would be useful to understand more about how to approach women regarding self sampling. However higher uptake was reported when sampling kits were directly mailed to women compared to an offer to collect or order a kit
- the potential for a negative impact on usual responders had not been explored.

Comments

The UK NSC would welcome comments on this proposal.

Forthcoming work

Data from the HPV pilot sites was presented to the UK NSC during its discussions on the above issues. This is currently being prepared for publication.

The UK NSC is in the process of initiating work to consider the options for screening in the vaccinated population. This will provide an opportunity to return to a number of issues and to take account of more recent data. An example of this is genotyping which, at the moment, is not being proposed as part of the primary screening strategy or as part of the surveillance strategy.

We will engage with stakeholders as this work develops.

Report to the National Screening Committee

Professor HC Kitchener, Chair Advisory Committee for Cervical Screening

June 2015

Contents

1. Summary

Cervical cancer is the second biggest cancer killer in women worldwide, but due to an effective cervical screening programme in the UK, deaths have fallen by around 60% to fewer than 1,000 deaths/year since the national programme was launched in 1988. This introduced structure, standards, quality assurance and a computerised system which is still in use. Up until now, screening has been based on cytology, but the causative role of high risk human papillomavirus, together with the introduction of the HPV prophylactic vaccination programme constitute a powerful rationale for switching secondary prevention to primary HPV screening and reserving cytology for HPV positive women. This is supported by four large European randomised trials, from which a subsequent pooled analysis has confirmed that, compared with cytology, HPV screening reduces the risk of developing cervical cancer. One of these trials, the ARTISTIC Trial, was conducted in England. What primary HPV screening offers is the prospect of greater sensitivity, compared with cytology, for the detection of underlying disease, and for those women who test HPV negative, longer protection allowing extension of screening intervals from 3 years to 6 years between ages 25 to 49, and potentially from 5 years to 10 years for women aged 50 and over. An exit HPV negative result for women aged over 60 is likely to confer protection from cervical cancer into an older age than is currently the case for women regularly screened by cytology. The challenge for primary HPV screening, however, is the management of women who test HPV positive/cytology negative, amongst whom the risk of developing cervical intraepithelial neoplasia grade 3 or worse (CIN3+) is twice that of the general population. This requires early recall at 12-24 months and referral if the infection persists. A detailed cost effectiveness modelling exercise based on ARTISTIC data and other UK data concluded that HPV as a primary screen would be cost and life years saving, but that triggering referral of these HPV positive/cytology negative women at 12 months would be more cost-effective than 24 months.

In order to confirm the performance of primary HPV screening in the NHSCSP, and to address the challenge of HPV positive/cytology negative screens, as well as ensuring general acceptability amongst women and primary care, and other practical issues, a large national pilot study, at 'Sentinel Sites' has been underway since April 2013. The pilot involves the entire screening age range (25-64) and involves HPV positive women having reflex cytology on the LBC sample. Abnormal cytology triggers referral to colposcopy and negative cytology requires early recall initially at 12 months (see Appendix). A 'safety check' requires routine recall at 3 years to ensure prior to moving to the anticipated extension of the screening interval to 6 years the detection of CIN3+ is sufficiently low. Laboratories have only partially converted to primary HPV, allowing both maintenance of cytology skills and contemporaneous comparison between primary HPV and primary cytology. These data are being evaluated independently and an interim report has been provided separately to the NSC. Initial data indicated that HPV followed by cytology had not increased baseline referrals when compared with cytology triaged by HPV. Early data censored in July/August indicated that 8.5% of screened women were HPV positive/cytology negative, of whom over 50% had cleared HPV at 12 months. Detection rates of CIN grade 2 or worse and CIN3 were significantly higher amongst women who had primary HPV screening compared with primary cytology. Larger numbers with updated data will permit more reliable interim findings. The pilot appears to have been well accepted by women and primary care, due in part to careful preparation and previous experience with HPV triage.

A national switchover to primary HPV screening would require a reconfiguration of laboratories both to maximise economy of scale with tests that have high throughput platforms and to maintain adequate cytology workload in a smaller number of laboratories; overall cytology work would fall by over 80%. Staff redundancy would be an issue though cytoscreeners are skewed towards older age and some could be redeployed in HPV testing.

A final but critical point is the Exeter computerised system which is over 20 years old would not be fit for a HPV based programme, and would need to be replaced by a new system. For safety reasons the Sentinel Sites pilot could not extend without a new system, and a new system for the country should be piloted within the Sentinel Sites project before full roll out.

2. Importance of cervical screening

Worldwide, cervical cancer is thought to be responsible for around 275,000 deaths per year which ranks it as the second most common cause of cancer deaths in women after breast cancer. Screening for pre invasive changes in cervical epithelium, by means of cervical cytology, has resulted in a major fall in both incidence and deaths in the developed world, but in developing countries, most women have neither access to screening nor access to adequate treatment. As a result of screening, cervical cancer now ranks 15th in female cancer deaths in the UK and mortality has fallen from 6.4 per 100,000 population in 1988 to 2.2 per 100,000 population in 2012[1].

3. Current screening strategy

3.1. Liquid based cytology

The mainstay of screening since the inception of the National Programme in 1988 has been exfoliative cytology. This allows detection of abnormal epithelial changes which triggers referral for colposcopy, biopsy and treatment. From 2006, the Cervical Screening Programme converted from conventional to liquid based cytology which allows reflex triage testing without a second sample being needed. Treatment for the premalignant lesion, cervical intraepithelial neoplasia (CIN) grades 2 and 3, is a straightforward outpatient procedure which effects cure in 90-95% of cases by means of a single excisional procedure. Treatment failure is usually detected in follow up, and a second treatment provided, but cancer will occur despite treatment in around 1 in 200 cases[2], which actually corresponds to around five times the population risk. In recent years, there has been concern that loop excision, which is the main technique employed to treat CIN, is responsible for an increased risk in preterm labour. In a recently published case control report from England, the authors found that increased risk of preterm and very preterm labour only really applied to excisions greater than 1.5cm deep[3], which in fact accounts for a small minority of treatments.

3.2. HPV testing

The causative role of high risk human papillomavirus (HR HPV) which was confirmed in the early 1990's has provided a precise means of primary prevention through prophylactic vaccination, and more precise strategy for secondary prevention through HPV testing in cervical screening.

HPV DNA testing has been developed because HR HPV status stratifies risk which can be exploited to triage management. Essentially, being HR HPV negative places a woman at very low risk, lower than negative cytology, whereas a HR HPV positive test can select for appropriate onward investigation or early recall. HPV testing is already in place in the screening programmes in England and Northern Ireland to select colposcopy referral (triage) amongst women with low grade cytological abnormalities, and also throughout the UK as a test of cure for women treated for CIN. This approach has had the principal benefit of accelerating return to routine recall, and advancing the diagnosis of underlying CIN because of avoiding the need for repeated cytology.

4. Purpose of proposed change

The rationale of the proposed change whereby HR HPV testing would replace cytology as the primary screen is based on greater sensitivity to detect CIN, extension of screening intervals, and now that the vaccinated cohort is coming through, a more precise means of detecting women at risk of disease.

5. Epidemiology

HR HPV infection occurs rapidly after sexual 'debut', with around 50% of females acquiring infection of the cervix within six months. Thereafter the immune system will clear the infection in the majority of cases, but type specific viral persistence is seen in around 30% of infections after 2 years[4], and it is this persistent infection by high risk types that is responsible for cellular changes resulting in CIN. The prevalence of infection is very age dependent, falling from around 40% at age 20-24 to around 7% at age 50[5]. Some studies have shown a small increase in prevalence in the menopausal age range. The mean age of acquiring HPV infection and the mean age of developing CIN3 (around age 30) suggest that the process of acquiring the true cancer precursor lesion generally takes 10-15 years. Low grade CIN (CIN1) is probably best regarded simply as chronic infection and is not treated, whereas CIN3 is thought of as the cancer precursor lesion and must be treated. CIN grade 2, which is usually associated with either grade 1 or grade 3, is also treated. While there has been a trend to manage low grade CIN more conservatively, treatment of CIN grade 2 or worse (CIN2+) remains the standard of care. CIN grades 1 and 2 are regarded as potentially regressive, and though some CIN3 may regress in young women, there is no biomarker to distinguish such lesions. The concept of a transition from CIN1 through CIN2 and CIN3 is an oversimplification, and many CIN3 lesions probably arise de novo. The change made in England in 2003 to the age threshold for initiating screening from age 20 years to 25 years, was based on the demonstration that screening was ineffective in preventing cancer in that age range, and that screening would detect a very large number of low grade lesions, many of which would regress risking unnecessary treatment[6]. Due to public pressure following a small number of deaths in women under 25 years, the decision to begin screening at 25 in England was reviewed in 2009 by the Advisory Committee for Cervical Screening[7]. This has subsequently been supported by the National Screening Committee, and is also supported by IARC.

The risk of CIN3+ developing in adult women found to have a HR-HPV infection, can be determined from ARTISTIC trial data[8]. Compared with HPV negative women whose risk of being found to have

CIN3+ over a six year period was 0.28%, baseline HPV 16 infection was associated with a 100 fold increased risk. For 'any type' of HR HPV positivity, it was 20 times greater. Even amongst women with negative cytology at baseline, there was an almost 10 times greater risk for women who were HR HPV positive compared with HR HPV negative.

6. Primary prevention

The bivalent and quadrivalent HPV vaccines have been shown in pivotal randomised trials to have 98% efficacy in preventing CIN2+ related to HPV types 16/18 amongst females who were HR HPV negative at the time of vaccination. In addition, it has been recognised that there is some cross protection from other types, which provides an added effect when compared with types 16/18 alone.

Based on trial efficacy data and genotyping data from a large cross section of cervical lesions in England, vaccination would be expected to prevent at least 70% of CIN3[9]. Vaccination will also have a large impact on the incidence of HR HPV infection, which will initially be seen in the 'catch up' campaign cohort of 14-18 year olds who were vaccinated between 2008-10. Reduced prevalence of HR-HPV should be detectable in England amongst 25 year olds being screened from 2015 onwards.

Between 2008-13 the bivalent vaccine was used in the UK before switching over to the quadrivalent vaccine in 2013. Based on the published bivalent vaccine efficacy[10] for types 16/18 and 31/33/45/52/58, genotyping data from ARTISTIC[4] suggests that vaccination could reduce the prevalence of HR HPV by around 50%. A recent report from Scotland, where screening still begins aged 20 years, has shown a significant reduction in CIN3 (RR 0.45; 95% CI 0.35-0.58)[11] amongst the vaccinated cohort. In a screening programme where HPV status determines the number of women requiring any further action, the expected impact of vaccination would therefore be considerable in terms of the proportion requiring reflex cytology, referral to colposcopy and treatment for high grade CIN.

7. The HPV test

Until recently, Hybrid Capture II^{R} (HC2)[12] was the only approved test, but there are now a number of CE marked tests, which have been shown in prospective testing in the triage setting within the screening programme to be as sensitive and more specific in the detection of CIN2+ than HC2[13]. Most of these tests are based on the detection of viral DNA; one detects viral RNA. Furthermore some of these new tests can exploit high throughput platforms developed for other microbiology diagnostic tests. HPV testing therefore offers the potential to concentrate testing into a smaller number of hubs maximising the economic benefits of high throughput.

8. The evidence supporting primary HPV screening

HPV primary cervical screening has been evaluated in four large European trials which involved at least two rounds of screening. These were conducted in the Netherlands (POBASCAM)[14], Sweden (SWEDSCREEN)[15], Italy (NTCC)[16], and England (ARTISTIC)[17]. All except ARTISTIC, which used liquid

based cytology, were based on conventional cytology, and all of these trials compared cytology with cytology plus HPV testing, with interventions for cytology negative/HPV positive women. In all four trials the HPV arm showed a reduction in the detection of CIN2+ in the second screening round, as a result of greater sensitivity achieved by HPV testing in round 1. Crucially for screening, a recently published pooled analysis of all four trials involving over 176,000 women with a median of 6.5 years follow up, showed clear evidence of a reduction in the incidence of cancer in the HPV arms compared with cytology alone; the hazard ratio for developing cancer was 0.6[18].

Some of the data used in this document have been drawn from the ARTISTIC Trial of HPV primary screening which generated the largest prospective genotyped dataset developed in the UK, and allows correlation of baseline screening results with clinical outcomes over a six year period. Not only does screening for HPV provide greater protection, but it also allows screening intervals to be extended. Data from ARTISTIC over three screening rounds and a mean follow up of 72 months, indicated that the cumulative rate of CIN2+ was similar after two rounds (3 year interval) following a negative cytology result as after three rounds (6 years) following a negative HR HPV test (0.73 vs 0.87). The cumulative rate of CIN2+ over a mean of 6 years, was 1.41% (1.19-1.65) for negative cytology at baseline compared with 0.87% (0.70-1.06) over 6 years for negative HPV. The corresponding data for CIN3+ was 0.63 (95% CI 0.48-0.80) for negative cytology compared with 0.28 (95% CI 0.18-0.40) for a negative HR HPV test[8]. For HPV negative women over 50 years, the cumulative risk over six years was only 0.16% (95% CI 0.07-0.34), suggesting the potential to extend the screening interval for women over 50 to 10 years.

Although the randomised trials of HPV testing involved co-testing with cytology, there is clear evidence from the ARTISTIC trial that co-testing (cytology and HPV) would not be cost-effective compared with HPV alone. There were 20,697 HPV negative women at baseline amongst whom 1497 and 46 were found to have low and high grade cytological abnormalities respectively. Amongst these, 9 CIN3 and 28 CIN2 lesions were identified in the first screening round, and 2 CIN3 and 2 CIN2 lesions in the second round. This means that co-testing would have required 20,000 additional cytology and up to 1500 colposcopies to detect 11 CIN3 lesions (PPV<1%). Therefore HPV negative women in whom abnormal cytology was identified were at low risk with cumulative rates over six years for CIN2+ and CIN3+ of 3.24% (95% CI 2.32-4.28) and 0.83% (95% CI 0.4-1.52) respectively. Indeed, the corresponding rate for the entire ARTISTIC population was not lower for CIN2+; 3.88% (95% CI 3.59-4.17) and was in fact lower for CIN3+; 1.96% (95% CI 1.76-2.17)[4].

9. The clinical performance of HPV testing

In event of a conversion from cytology to HPV as the primary screen, it is pertinent to consider differences in clinical performance. Currently, negative cytology means a return to routine recall every three years up until age 50, and every 5 years between 50-64 years. In England and Northern Ireland, a borderline or low grade dyskaryosis involves reflex testing for HR HPV, so called triage. This allows the 30-40% who are HR-HPV negative to be returned to recall whereas women who are positive are referred immediately to colposcopy; this had a positive predictive value (PPV) in the triage pilot for diagnosing CIN2+ of around 16%[19]. Women with high grade cytological abnormalities are referred directly to colposcopy with a PPV of 75-90%[20].

In the event of conversion, HPV testing would enable an extended screening interval of six years if negative. If positive, reflex cytology would be performed and any grade of abnormality would result in referral. The PPV would therefore remain the same for the combination of HPV positive/low grade cytology. As almost all high grade cytology is HPV positive, the PPV would be expected to remain unchanged at 75-90%. The challenge for primary HPV screening is that because of the prevalence of HPV infection, particularly in younger women, it is less specific in terms of underlying CIN grade 2 or worse, than cytology. HPV testing with cytology triage will therefore produce a new class of abnormal result; those who are HR HPV positive/cytology negative, which was found in around 9% of the ARTISTIC cohort who were aged 20-64. This proportion would be expected to be lower (~8%) in a population aged 25-64, and of course far lower in a vaccinated population. Although this group are at twice the risk of the general population, referral would await evidence of persistent infection as many women clear the infection over 12-24 months. There are potential biomarkers to improve the specificity with respect to underlying CIN2+, and one of these which combine p16 and Ki67 is currently under evaluation in the Sentinel Sites primary HPV pilot project. There is no reason to believe that lesions detected by HPV triaged by cytology should be different from those detected via cytology triaged by HPV.

9.1. What is the comparative accuracy of HPV DNA tests and cytology for: CIN2 or worse (CIN2+) and CIN3 or worse (CIN3+), in women under 30 years old?

The comparative accuracy of primary HPV with cytology triage can be expected to be less specific but more sensitive than primary cytology with HPV triage, because of the high prevalence of HPV in women 25-29. Therefore the rate of referral at baseline could be expected to be higher than following primary cytology in this age group. Although the PPV for CIN2+ may be lower amongst those referred following primary HPV compared with primary cytology, the detection rate overall for CIN2+ and CIN3+ should be greater than for primary cytology in the 25-29 age group. Crucially, cytology negative/HPV positive women will harbour undetected CIN, amongst whom at least 3% will have CIN2+, which will be detected as a result of early recall. Data will be available in the report of the primary HPV pilot.

9.2. What is the accuracy of HPV DNA tests compared with liquid based cytology for these outcomes in all women and in women under 30 years old?

With respect to women below 30 and older than 30, tests should perform similarly from a purely analytical point of view. Despite the greater prevalence of HR-HPV infection, the greater prevalence of CIN2+ and CIN3+ in women under 30, should mean that the PPV for the detection of high grade CIN amongst those referred at baseline could be expected to be similar to the age group 30 and older. The proportion who are HPV positive/cytology negative will be greater in the under 30's which would therefore result in a disproportionately large number of the HPV positive/cytology negative women under 30 who require subsequent referral to colposcopy, either because of persistent infection or abnormal cytology. Around 8% of screened women are in the category of HPV positive/cytology negative. If 80% adhere to early recall and 40% of these show persistent HPV over 12-24 months, then an additional 2.5% of screened women could be referred to colposcopy, based on current data. It must be borne in mind however, that this would include a prevalence effect which will lessen for that cohort in subsequent rounds. It should also be noted that amongst 25 year olds, the impact of HPV 'catch up' vaccination programme will be felt beyond 2015, and the vaccination of 12/13 year olds beyond 2020, such that the prevalence of high risk HPV infection will

be markedly reduced. Any increase in the requirement for colposcopy should therefore be temporary. Again, data comparing HPV testing in the 25-29 and 30 and older age groups will appear in the pilot report.

9.3. What is the estimated rate of over-diagnosis of regressive lesions when current practice and HPV DNA screening are compared?

The estimated rate of regressive lesions is not possible to determine with any precision because CIN2+ is treated and not managed expectantly. Around 25% of detected CIN is CIN grade 1, which is not treated. CIN grade 2 though treated, is not widely regarded as a true precursor lesion but there have been no randomised trials of treatment versus observation to determine regression rates. It should be noted that CIN2 is not a robust diagnosis; it is often reported as CIN1/2 and CIN2/3 as it often occurs amongst areas of CIN1 and CIN3. The WHO is moving towards the reporting of CIN as low grade and high, and CIN2+ would be categorised as high grade. CIN3 may regress in a small proportion of cases but non treatment of CIN3 would be regarded as clinically negligent, even though not every case of CIN3 would progress to cancer. The most regressive lesions are CIN grade 1 which some consider to be little more than evidence of a persistent HPV infection. Amongst those women who are HPV positive and have abnormal cytology, the profile of detected CIN should be similar to primary cytology and HPV triage. It is possible that amongst women with negative cytology and persistent HPV, more CIN1 will be reported at the time of early recall.

9.4. What proportion of women with HPV positive results will be cytology negative? Can this be broken down by under 30s and over 30s?

As stated above, around 8% of screened women aged 25-64 might be expected to be HPV positive/cytology negative. In the ARTISTIC trial this proportion was 9% overall; 16% in the 20-29 and 2.6% in the 30-64 range. In the age range 25-29, the proportion is likely to be 11-12% as the HPV prevalence dropped from 37.5% to 27.5% in the 20-24 and 25-29 age ranges respectively. More precise data will be available from the report of the primary HPV pilot.

9.5. Have cut-offs for HPV testing been agreed and has the frequency of screening been agreed?

HPV tests are analysed according to manufacturer's instructions. The Hybrid Capture 2 (HC2) test did allow a variable cut off to be used and a cut off of 2pg/Co was found to be clinically more useful than the manufacturers cut off of 1pg/Co in terms of a beneficial balance of sensitivity and specificity. The recently developed tests which laboratories generally prefer for reasons of high throughput automated platforms being used for other microbiology tests, are used according to manufacturer's instructions. As stated earlier, these newer commercial tests compared favourably with HC2 at a cut-off of 2pg/Co, when tested in the triage setting. To compare the tests robustly in the primary screening setting would require very large expensive studies which cannot be justified. As stated in Section 8, there are prospective data from ARTISTIC which indicate that six yearly screening until age 50, and 10 yearly screening over 50, would be as effective as three and five yearly screening as currently used for cytology. The conversion of the Australian programme in 2016, from 2 yearly primary cytology to primary HPV screening, envisages 5 yearly HPV screening.

9.6. What is the proposed diagnostic pathway for HPV positive women?

In general, HPV positive samples are subjected to reflex cytology triage in the liquid residue, with immediate colposcopy referral for abnormal cytology. Those who are cytology negative are at twice the risk of having or developing CIN grade 2 or worse over six years compared with the general population and therefore harbour disease not detected cytologically. Early recall is therefore appropriate. In the pooled analysis of the RCT's of HPV primary screening, the incidence of cervical cancer following primary cytology began to increase relative to HPV primary screening after 2-3 years, so while there is no urgency for early recall in terms of preventing cancer, there is a balance to be struck between an interval long enough to maximise viral clearance, but not so long that women may not adhere to recall. The current pilot protocol involves early recall at 12 months and if HPV persists, again at 24 months prior to colposcopy referral. Several sites however have started to refer women who are persistently types 16/18 positive at 12 months. This is because some tests offer a 16/18 readout, and the specificity of HPV positivity can be increased by restricting referral to the highest risk types in terms of disease. If employed immediately, too many young women would be referred, on the other hand recall at 12 months for 16/18 positive and further recall at 24 months for other high risk positives will allow the highest risk women to be colposcoped and allow further clearance to occur in those with lower risk types. National programmes will vary according to follow up protocol but the economic modelling recently reported from the ARTISTIC group[21], suggested that selective referral to colposcopy at 12 months was found to be more cost effective than delaying referral for all until 24 months. This is probably related to the model predicting increased non adherence to repeated early recall. Although HPV positive/cytology negative women do not require immediate intervention, they do require early recall. To what extent this risk and need to recall is understood by women is not clearly known, however early experience from the English pilot study has indicated that adherence to early recall is encouraging, and qualitative research to address the views and experience of screened women is planned for 2015.

9.7. Have evidence based policies for the diagnostic pathway been agreed and published?

The pathway employed (see Appendix) in the national pilot broadly reflects the principal considerations above, i.e. referral with abnormal cytology and early recall with negative cytology. Of the six centres, some refer if 16/18 positive at 12 months, and other continue to defer referral until HPV persists until 24 months. This comparison will be of interest.

9.8. The treatment

This will remain unchanged. CIN1 is not treated and CIN grades 2 and 3 are treated. This is in line with worldwide practice.

9.9. Is there RCT evidence that the screening programme is effective in reducing mortality or morbidity?

As stated on page 5, the pooled analysis of the four European RCT's of primary HPV screening indicated that HPV significantly reduced the incidence of cancer. Overall the rate ratio was 0.60 (95% CI 0.40-0.89), and for negative baseline screening it was 0.30 (95% CI 0.15-0.60). There are not yet data on mortality. Reduced incidence of cancer will be cost saving and reduce cancer treatment related morbidity.

9.10. When the test produces information upon which no immediate intervention is required (eg HPV positive/cytology negative) is there evidence that this accurately measures risk, is valued and easily understood by the recipients?

The risk is accurately known from 6 year follow up data in the ARTISTIC trial. HPV positive/cytology negative women had a cumulative incidence of CIN2+ over two subsequent rounds (at three and six years) of 3.6% (1.87 + 1.73) whereas the corresponding figure for all women who screen cytology negative was 1.29% (0.42 + 0.87). We do not yet have data to show to what extent this will be appreciated by women, but the encouraging adherence to early recall in the pilot suggests it is, by the majority of women.

9.11. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).

The principal harm that could accrue from HPV primary screening could be a) increased anxiety as a result of being HPV positive/cytology negative thus necessitating early recall in place of return to routine recall, b) increased morbidity from increased detection of CIN2+, some of which would not result in cancer. With respect to the latter, the pooled analysis of RCT's suggests that the overall effect is a fall in cancer incidence amongst women screened with HPV.

In general, the process of cervical screening is viewed favourably as a public health measure, with the benefits of screening, i.e. the reduction in cervical cancer deaths outweighing the harm. It should be noted that in the longer term, the observed effects of vaccination should result in a large reduction in the incidence of high grade CIN and consequently, reduced treatment-associated morbidity. The theoretical harm that could occur from the knowledge of having an oncogenic, sexually transmitted infection has not been apparent. Throughout not only the primary HPV pilot, but also the preceding pilot of HPV triage, there has not been evidence of problems either from women, GP's or sample takers. The care that has been taken to inform women, as well as sample takers and doctors has undoubtedly help to avoid distress.

10. How the screening programme would change with HPV primary screening

Given the evidence from the European trials described above, HPV based screening would be expected to save additional lives compared with current provision. If the hazard ratio from the trials were translated into population based screening there would be a significant reduction in cancer incidence with a similar reduction in deaths. With the vaccination programme in place however, primary prevention is already predicted to save a proportion of these lives; nonetheless a more sensitive HPV based programme would more precisely target women at risk and prevent cancers and deaths. It would also be more efficient and cost effective.

There would be two major changes in a HPV based programme. The first would be a reduction in the number of screens in a lifetime. Currently a woman would receive 12-13 invitations, (3 yearly intervals until age 50 and 5 yearly thereafter) but with HPV testing, this could be essentially halved due to extending screening intervals to 6 yearly and 10 yearly over the age of 50 years, based on extended follow up data from ARTISTIC. Additionally, as the vaccinated population grows older, together with the prospect of a protective nonavalent vaccine[22], there is the future prospect of de-intensifying screening further according to the evidence emerging from post vaccination surveillance

programmes. It should also be noted that an 'exit' screen at 64 years by means of HPV testing will provide longer protection into older age compared with that of cytology, which has been estimated on the basis of having been regularly screened until 64, to continue until the mid 70's.

The second characteristic of a HPV based programme that will differ is the creation of the already highlighted HPV positive/cytology negative cohort, who based on ARTISTIC data are at twice the population based risk of developing CIN2+. As stated above, it will be necessary to institute early recall of these women to maximise the benefit of the greater sensitivity of HPV testing without swamping colposcopy referral and creating unnecessary anxiety. Early recall is a key component in the national pilot.

11. Acceptability of HPV screening

There has been awareness since HPV testing was first piloted in triage and test of cure, that using a test which detects an oncogenic sexually transmitted virus could present difficulties for women. As already stated, it has been noticeable how well HPV testing appears to have been received by women and health professionals, as there has been an absence of adverse comment from women, GP's and sample takers. Throughout the piloting and implementation of HPV testing there have been no formal complaints from any quarter about the consequences of HPV testing. That is not to say that individual women have not been concerned about what a positive HPV result means to them personally, but the careful planning of information provision and training of sample takers appears to have been successful in transferring from the trusted 'smear test' to liquid based samples which have incorporated both cytology and HPV test results. A formal evaluation of the acceptability of primary HPV screened is currently planned within the primary HPV screening pilot.

12. Cost effectiveness of HPV primary screening

A number of cost effective evaluations of primary HPV screening has been performed, in the Netherlands[23], Germany[24], Norway[25] and Canada[26], all of which have determined that changing from cytology to HPV based primary screening would be favourable. International cost effective analyses may lack country specific considerations such as screening and vaccine coverage, compliance with follow-up, likely algorithms and validation of a modelling platform fitted to country specific data. In a recently published[21] cost effectiveness analysis of primary HPV screening in England, these considerations were all accounted for in a detailed analysis which concluded that switching from liquid based cytology to primary HPV screening would be both cost saving and more effective across a number of strategies, for both unvaccinated and vaccinated cohorts. The methodology for this study involved a model platform with three main components; a dynamic model of sexual behaviour and HPV transmission, a Markov cohort model of the natural history of CIN and invasive cervical cancer, and a cohort/multi-cohort model of screening, diagnosis, treatment and follow up. The model platform which had been used previously to evaluate changes to screening in Australia and the UK[27, 28], was validated against data from the ARTISTIC Trial which represents generalisable screening results in England. A number of necessary assumptions regarding sexual behaviour, vaccination data and other screening data were drawn from UK screening data sets, NATSAL II, and UK vaccination data. Because of uncertainty regarding likely final screening algorithm, particularly

with respect to women who were HPV positive/cytology negative, four screening strategies were studied.

1. HPV primary screening with a 24 month colposcopy for women with negative cytology and persistently positive HPV testing.
2. As for 1., but referring for colposcopy at 12 months on the basis of persistent HPV type 16/18.
3. Referring for colposcopy at baseline if women were positive for HPV type 16/18.
4. Screening with both HPV and cytology and referring either because of high grade cytology or managing women who are HPV positive with negative cytology, as in strategy 2.

Overall, cost savings compared with current practice, i.e. liquid based cytology, were predicted for both vaccinated and unvaccinated cohorts. These cost savings ranged from 9% (strategy 4) to 22% (strategy 1) in vaccinated cohorts and from 7% (strategy 4) to 18% (strategy 1) in unvaccinated cohorts. The most effective strategy (strategy 3) involved direct referral to colposcopy of women who screened HPV16/18 positive, however this would involve a considerable increase in colposcopy which was not considered feasible. This suggests however, that using cytology to triage HPV positive women, and referring on the basis of persistent HPV16/18 at 12 months may be more feasible and remain effective.

In general strategies using HPV as a sole primary screening test were both cost saving and life years saving. It was also determined that using 12 months rather than 24 months referral to colposcopy for persistent negative cytology/HPV positive, was more effective. Having a 12 month follow up prior to colposcopy referral resulted in an estimated 73-113 and 37-41 additional life years saved per 100,000 women in unvaccinated and vaccinated cohorts respectively, when compared with a 24 month referral. This difference is presumably due to the likelihood that more women would fail to adhere to follow up over 24 months compared with 12 months. Exploratory analyses also predicted that retaining cytology only below ages 30 or 35 years would increase costs and reduce effectiveness. This suggests that implementation of HPV as a sole screen from aged 25 years would be optimal in terms of cost and life years saving in England. In conclusion the modelled analysis predicted that primary HPV screening would be both more effective and cost saving compared with current practice. Adhering to early recall for women who test HPV positive/cytology negative is seen as of key importance if the sensitivity of HPV over cytology is not to be undermined. The analysis supports a switch from cytology to HPV based cervical screening.

13.English primary HPV screening pilot

The large pilot study of primary HPV screening, already referred to, was initiated in April 2013 across the six Sentinel Sites used previously to pilot triage and test of cure. It has been designed to assess; feasibility and practicability, to test clinical algorithms, to support an economic evaluation and assess acceptability by women. Interim data have been reported in a separate report to the NSC.

14. Consideration of infrastructure and phasing for national switchover from cytology to HPV

There are three key issues which need to be considered in the event of a decision to proceed to a national conversion from primary cytology to HPV.

1. Laboratory capacity and reconfiguration.
2. The computer system to support the changed programme.
3. The implication for staff given a massive reduction in cytology.

14.1. Laboratory capacity and reconfiguration

There are two principal drivers here. The first of these is the need to maximise cost effectiveness through high throughput testing using the capacity of the HPV testing platforms. The second, and in some respects the dominant driver, is the need to maintain laboratories with adequate workload for cytology. Current guidance supports a minimum laboratory throughput of 35,000 cytology slides per year to maintain expertise amongst 5-7 screening staff. Given that cytology would be reduced to around 15% of its current level based on the pilot experience, this would require around 200,000 HPV samples per year to generate 35,000 cytology slides. Given around 3 million screened women per year, this would equate to concentrating cervical screening to around 15 labs in England, perhaps two each in Scotland and one each for Wales and Northern Ireland. Whether this were configured as microbiology and cytology in separate but co-located labs, or HPV testing and cytology conducted in a single lab would be determined locally. Both models have been used in the pilot and both work well. Commissioning more centralised services will present some challenges.

14.2. The computer system

The current cervical screening programme uses a call/recall IT programme developed 25 years ago. Should HPV screening be introduced, the increasingly personalised screening intervals and varying results will require a modern IT system. The invitation of the cohort and management of women through a much more complex pathway without running the risk of losing women will be absolutely crucial. Failure to invite and absence of failsafes to ensure completion of the pathway (when the woman consents so to do) will result in women being lost to the system and getting cancer when they need not. An IT system that is not fit for purpose risks loss of life, massive reputational damage and loss of confidence in the programme.

14.3. Implications for staffing

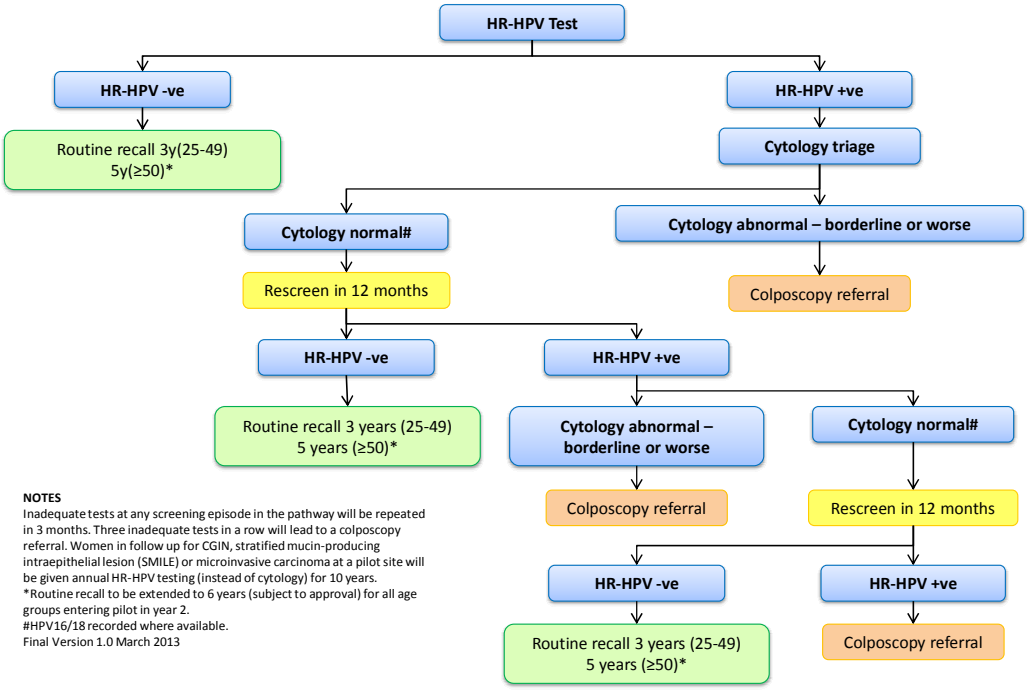
It is inevitable that a major reduction in cytology will require fewer staff, however a number of factors may mitigate this problem. These include the age range of the cytoscreeners which tends towards older staff, the potential to redeploy from cytology to HPV testing, and cytology staff seeking new posts when it becomes clear that redundancy threatens. With respect to the last point there is a need to try to clarify timescales to avoid a premature flight from cytology labs. Although laboratories have only partially converted in the pilot, once a decision to switch has been made, it would be better to plan for a total switchover in order to avoid a perceived two tier programme based partly on primary HPV and partly on primary cytology. There will be a need to consider how to mitigate a sudden move from 3 yearly to 6 yearly screening intervals in order to avoid an interval during which activity collapses. SCHARR in Sheffield University have been tasked with looking at this. Their report will be available in Q2 of 2015.

15. Conclusion

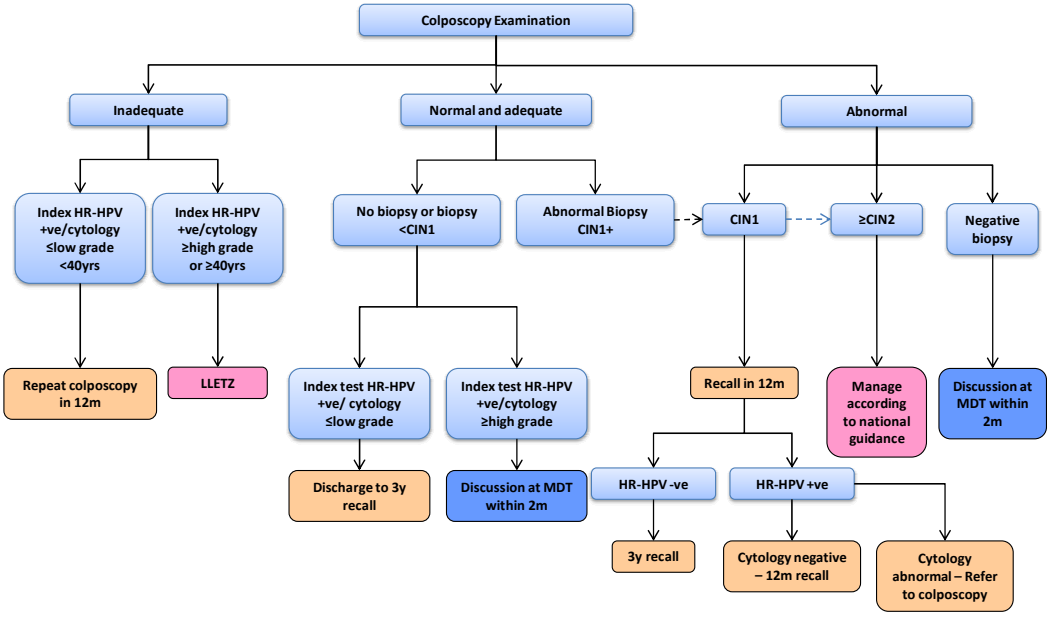
In conclusion, there is grade A evidence to support a switch from primary cytology to primary HPV testing in cervical screening. This should save life years and cost less, increasingly so as the vaccinated population grows older. A NHS pilot exercise to demonstrate feasibility, practicability and safety is reported separately. There will be some challenges in terms of laboratory reconfiguration, computer systems and staffing, but in an era of primary prevention through HPV vaccination, HPV based screening offers the prospect for a more effective, more streamlined, cheaper programme which is more tailored to individual risk.

Appendix

Algorithms for HPV primary cervical screening currently in use with the primary HPV screening pilot in England



HPV Primary Screening Algorithm – Pilot Year 1:
All women aged 25-64 on routine call/recall and early recall



HPV Primary Screening Pilot: Colposcopy Management Recommendations

References

1. CancerResearchUK. Cervical cancer mortality statistics. 2013. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/cervix/mortality/.
2. Soutter WP, de Barros Lopes A, Fletcher A, et al. Invasive cervical cancer after conservative therapy for cervical intraepithelial neoplasia. *Lancet* 1997; **349**(9057): 978-80.
3. Castanon A, Landy R, Brocklehurst P, et al. Risk of preterm delivery with increasing depth of excision for cervical intraepithelial neoplasia in England: nested case-control study. *BMJ* 2014; **349**: g6223.
4. Kitchener HC, Almonte M, Gilham C, et al. ARTISTIC: a randomised trial of human papillomavirus (HPV) testing in primary cervical screening. *Health Technol Assess* 2009; **13**(51): 1-150, iii-iv.
5. Kitchener HC, Almonte M, Wheeler P, et al. HPV testing in routine cervical screening: cross sectional data from the ARTISTIC trial. *Br J Cancer* 2006; **95**(1): 56-61.
6. Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. *Br J Cancer* 2003; **89**(1): 88-93.
7. AdvisoryCommitteeonCervicalScreening(ACCS). Extraordinary Meeting to re-examine current policy on cervical screening for women aged 20-24 years taking account of any new evidence and to make recommendations to the National Cancer Director and Ministers. 2009. http://www.cancerscreening.nhs.uk/cervical/cervical-review-minutes-20090519.pdf.
8. Kitchener HC, Gilham C, Sargent A, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *Eur J Cancer* 2011; **47**(6): 864-71.
9. Howell-Jones R, Bailey A, Beddows S, et al. Multi-site study of HPV type-specific prevalence in women with cervical cancer, intraepithelial neoplasia and normal cytology, in England. *Br J Cancer* 2010; **103**(2): 209-16.
10. Paavonen J, Naud P, Salmeron J, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. *Lancet* 2009; **374**(9686): 301-14.
11. Pollock KG, Kavanagh K, Potts A, et al. Reduction of low- and high-grade cervical abnormalities associated with high uptake of the HPV bivalent vaccine in Scotland. *Br J Cancer* 2014; **111**(9): 1824-30.
12. Qiagen, inventor Digene HC2 HPV DNA test. Crawley, UK.
13. Moss SM, Bailey A, Cubie H, et al. Comparison of the performance of HPV tests in women with abnormal cytology: results of a study within the NHS cervical screening programme. *Cytopathology : official journal of the British Society for Clinical Cytology* 2014.
14. Bulkmans NW, Rozendaal L, Snijders PJ, et al. POBASCAM, a population-based randomized controlled trial for implementation of high-risk HPV testing in cervical screening: design, methods and baseline data of 44,102 women. *Int J Cancer* 2004; **110**(1): 94-101.
15. Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 2007; **357**(16): 1589-97.
16. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010; **11**(3): 249-57.
17. Kitchener HC, Almonte M, Thomson C, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol* 2009; **10**(7): 672-82.
18. Ronco G, Dillner J, Elfstrom KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014; **383**(9916): 524-32.

19. Kelly RS, Patnick J, Kitchener HC, Moss SM. HPV testing as a triage for borderline or mild dyskaryosis on cervical cytology: results from the Sentinel Sites study. *Br J Cancer* 2011; **105**(7): 983-8.
20. ScreeningandImmunisationteamHealthandSocialCareInformationCentre. Cervical Screening Programme, England. Statistics for 2013-14, 2014.
21. Kitchener H, Canfell K, Gilham C, et al. The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England: extended follow-up of the ARTISTIC randomised trial cohort through three screening rounds. *Health Technol Assess* 2014; **18**(23).
22. Joura EA, Giuliano AR, Iversen OE, et al. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N Engl J Med* 2015; **372**(8): 711-23.
23. van Rosmalen J, de Kok IM, van Ballegooijen M. Cost-effectiveness of cervical cancer screening: cytology versus human papillomavirus DNA testing. *BJOG* 2012; **119**(6): 699-709.
24. Sroczynski G, Schnell-Inderst P, Muhlberger N, et al. Cost-effectiveness of primary HPV screening for cervical cancer in Germany--a decision analysis. *Eur J Cancer* 2011; **47**(11): 1633-46.
25. Burger EA, Ortendahl JD, Sy S, Kristiansen IS, Kim JJ. Cost-effectiveness of cervical cancer screening with primary human papillomavirus testing in Norway. *Br J Cancer* 2012; **106**(9): 1571-8.
26. Vijayaraghavan A, Efrusy MB, Mayrand MH, Santas CC, Goggin P. Cost-effectiveness of high-risk human papillomavirus testing for cervical cancer screening in Quebec, Canada. *Canadian journal of public health = Revue canadienne de sante publique* 2010; **101**(3): 220-5.
27. Canfell K, Sitas F, Beral V. Cervical cancer in Australia and the United Kingdom: comparison of screening policy and uptake, and cancer incidence and mortality. *The Medical journal of Australia* 2006; **185**(9): 482-6.
28. Creighton P, Lew JB, Clements M, et al. Cervical cancer screening in Australia: modelled evaluation of the impact of changing the recommended interval from two to three years. *BMC public health* 2010; **10**: 734.

Clinical impact and cost-effectiveness of primary human papilloma virus testing

Centre for Infectious Disease Surveillance and Control (CIDSC), Public Health England, London, UK

Executive Summary

A stochastic, individual-based model of human papillomavirus (HPV) infection and natural history leading to cervical cancer was used to evaluate the cost-effectiveness of primary HPV testing in the currently unvaccinated adult female population. The model was used to compare lifetime clinical and economic outcomes for screening cohorts following three protocols:

1. primary cytological screening followed by HPV testing of women with borderline or mild cytology results ("primary cytology protocol") with a recall interval for screen negatives for women of 3 years for women aged 25 to 50 years, and 5 years for women over 50 years;
2. primary testing for high-risk HPV followed by cytology of HR HPV positives ("primary HPV protocol") with the same recall interval as above;
3. primary testing for high-risk HPV followed by cytology of HR HPV positives ("5 year primary HPV protocol") with a 5 year recall interval for screen negatives women of all ages.

Compared to the primary cytology protocol, the standard recall primary HPV protocol, as modelled in this report, is expected to:

a. lead to a 4% increase in primary screens; an 18% increase in number of colposcopies; and a 29% increase in detection of cervical intraepithelial lesions of grade 2 or worse.
b. lead to a median decrease in cervical cancer incidence of 310 cases per year, and reduction in cancer-related deaths of 73 per year.
c. lead to a saving of 0.0026 discounted life years per women. The impact on quality adjusted life years (QALYs) is not well determined and is shown to be highly sensitive to the choice of screening-derived QALY detriments.
d. reduce net health-related costs by £15.8 million per year (due to the decrease in cytological testing, which is more expensive than HPV testing; in addition to a saving on cancer treatments).

Compared to the primary cytology protocol, the 5 year recall primary HPV protocol, as modelled in this report, is expected to:

a. lead to a 17% decrease in primary screens; while providing a 14% increase in detection of cervical intraepithelial lesions of grade 2 or worse.
b. lead to a median decrease in cervical cancer incidence of 159 cases per year, and reduction in cancer-related deaths of 54 per year.
c. lead to a saving of 0.0008 discounted life years per women. As above, the absolute impact on QALYs is not well determined, but the trade-off between screening- and cancer- related QALY losses means that a switch to a 5 year primary HPV protocol is more favourable in terms of net quality-adjusted life years than a switch to standard primary HPV testing.
d. reduce net health-related costs by £35 million per year

Introduction

Since the introduction of the National Health Service Cervical Screening Programme (NHSCSP) in England, in 1988, the primary assessment has been based on cytology testing to identify cervical abnormalities. Currently, testing for high risk human papillomavirus (HR-HPV) is used to determine management of women with borderline or low-grade abnormalities, and as a test-of-cure for recently treated women. In 2013, a pilot study of implementing screening in which the primary assessment is a test for HR-HPV was initiated at several sites across England.

The clinical evidence suggests that HPV testing has higher sensitivity for high-grade lesion detection[1,2], and provides stronger negative predictive power than cytology [3,4]. The costs associated with HR-HPV testing are also favourable compared to cytology testing. The purpose of this study is to evaluate the potential impact and cost-effectiveness of national implementation of HPV testing across the NHSCSP. To do this, we use a stochastic, individual-based simulation model that we have developed to characterise HPV infection and the natural history leading to cervical cancer. The framework builds on existing compartmental markov-models used to appraise UK vaccine policy [5] and screening practice [6]. This work is part of a longer-term project that will integrate this model with our existing transmission dynamic model of HPV vaccination in order to have a single model that can investigate the overall impact of combined vaccination and screening strategies.

Methods

Model

A stochastic, individual-based simulation model is used to evaluate primary HPV testing and the current primary cytology protocol. The key model components are: (a) acquisition of HPV infection; (b) natural progression of HPV infection, cervical intraepithelial neoplasia (CIN) and cervical cancer; and (c) detection and treatment of women with cervical abnormalities through cervical screening. Women are categorised according to HPV infection status, as illustrated in Figure 1. The model simulates a large population of women with individual histories. Women can acquire multiple, possibly simultaneous, HPV infections, and each infection follows its own timeline to clearance or emergence of a pre-invasive cancer lesion, adenocarcinoma or squamous cell carcinoma. Women undergo screening and the life history is changed according to any treatment undertaken (screening algorithms illustrated in Figures 2-3).

The risk of **HPV acquisition** is determined by a number of behavioural factors: (i) age of sexual debut; (ii) acquisition of new partners; (iii) duration of partnerships; (iv) frequency of sex acts; and (v) age of new partners. In the model, these behavioural components are parameterised using data collected by the National Survey of Sexual Attitudes and Lifestyles 2010 (NATSAL-3)[7]. We generated a model of sexual behaviour that captures decrease in sexual activity with age, as well as heterogeneity among individuals of a given age (described in more detail in the appendix A1).

A static model of **transmission** was applied in which male prevalence was assumed to be constant throughout the duration of model simulation; i.e. the introduction of primary HPV testing in cervical screening is assumed to have no effect on the prevalence of HPV in males. The probability of transmission of HPV is described as a function of (i) HPV prevalence among male partners according to

age; and (ii) the probability of transmission per contact with an infected individual (described in more detail in the appendix A2). The rate of HPV **clearance** is modelled by a decreasing function of time post-infection using a weibull distribution. The prevalence of HR-HPV among the female English population has been well characterised[8]. The model was calibrated using pre-vaccination surveillance data collected by PHE that measures type-specific prevalence of HPV in women[9,10] and HPV sero-prevalence measured in males[11] (described in more detail in the appendix A2). A MCMC algorithm was implemented in R to simultaneously identify the posterior distribution for the probability of transmission, clearance and male sero-conversion for each model HPV strain. As a validation of the parameterisation process, we compare the HPV positivity expected by the model under a primary HPV testing protocol, and the HPV positivity observed between May 2013 and August 2014 in the primary HPV pilot study (Figure 4). The model is parameterised completely independently of the primary pilot dataset, however, we are satisfied that the observations lie within the 95% prediction interval.

Disease progression and regression are modelled as continuous processes; the probability of a given cytological abnormality is determined as a function of time since infection. We use a nested conditional probability structure to generate a model in which the probability of a normal outcome decreases, while the probability of a severe outcome increase with time since infection (described in more detail in the appendix A3). The model was calibrated using observed cytological outcome and HPV typing data measured as a function of age in residual samples collected by the NHSCSP[10]. Incidence of invasive squamous cell carcinoma and adenocarcinoma of the cervix was characterised by the increasing risk of disease progression as a function of time following high-risk HPV infection, using a gamma distribution to model the wait time to a squamous cell carcinoma or adenocarcinoma. Cancer incidence was calibrated using cancer registrations in England reported by ONS and evidence from the NHSCP audit of cervical cancers[12]. A MCMC algorithm was implemented in R to simultaneously identify the parameters defining the natural progression of cytological abnormalities to cervical cancer for each model HPV strain, in a population that is undergoing screening according to the current national algorithm. The parameterisation is described in full in appendix A3.

The **screening behaviour** of women is characterised using age-dependent attendance as reported by the cervical screening programme and lifetime behavioural screening patterns derived from data collected by the cervical cancer audit team (personal communication with Alex Castanon & Peter Sasieni). The age at first screen is well characterised by a 'delayed' lognormal distribution. The waiting time to subsequent screens, under a standard recall, is modelled as a function of previous 'punctuality' (described in detail in appendix A4). This framework captures the behaviour of women who regularly attend screening appointments within a small window of their recall date; women who consistently demonstrate poor adherence to the recommended screening appointments; and women who begin with a poor adherence record but then switch to regular screening adherence behaviour.

Screening

Two alternative strategies were considered: (i) primary cytological screening with HPV testing to determine further management of cytology abnormals ("primary cytology protocol"), which is current screening practise, and (ii) primary HPV testing ("primary HPV protocol"), with cytology testing to determine further management of HR HPV positives.

Under the primary cytology protocol, a negative test leads to recall in 3 years (or 5 years for women over 50 years old); a high grade cytological outcome leads directly to a colposcopy referral; and identification of a borderline or mild cytological abnormality is followed by HPV triage where a negative HPV outcome leads to a standard recall, while a positive result leads to an immediate colposcopy referral (Figure 2). Under the primary HPV protocol, a negative HR HPV test leads to recall to screening in 3 years (or 5 years for women over 50 years old), while a positive HR HPV test results lead to cytological assessment of the same sample; all non-negative cytological results (including borderline) are referred to colposcopy; a negative cytology leads to a 12-month follow up. In the follow up arm, 3 successive positive HR HPV results lead to referral for colposcopy (Figure 3).

The actions following colposcopy are the same in both protocols. A negative outcome at colposcopy is assumed to lead to discharge to standard recall; CIN1 is untreated but leads to a 12 month follow up; while identification of precancerous lesions of grade CIN2 or worse leads to treatment followed by 'Test of Cure' triage at 6 months.

The sensitivity of cytological testing is explicitly built into the model; cytology outcome is defined probabilistically and varies as a function of time since infection (in detail in appendix A3). The sensitivity of the HPV test was assumed to lie between 90-95% for high risk HPV.

Attendance and outcome at colposcopy under a primary cytology protocol are constrained according to cytology result at referral, as reported by the cervical screening programme 2012-2013 (Table 1). The probability of attending colposcopy, and the likely outcomes, are assumed to be identical for women referred following low-grade cytology followed by HPV positivity under a primary cytology protocol, as for women referred for a positive HPV test followed by low-grade cytology result under a primary HPV protocol[13]. Colposcopy outcomes for women referred following a positive HPV test and high-grade cytology, under primary HPV protocol, are not significantly different from those reported following a high grade referral under the current primary cytology protocol. This has been evidenced in preliminary data from the pilot primary HPV programme (Table 1).

We assumed that of all cases of CIN2 or worse that should all be recommended for treatment, 83.1% return for treatment and 66.0% attend follow up appointments (source: cervical screening programme 2012-2013). The split between diagnostic biopsy and excision for those women that undergo treatment was assumed to be 63.2:2.6 in those originally referred due to low grade abnormalities, and 37.6:49.1 in those attending colposcopy following a high grade referral (source: cervical screening programme 2012-2013). In the absence of recent data to inform this model parameter, the type of procedure recommended is assumed to be unchanged in the context of the HPV primary screening, however, this decision may be sensitive to knowledge that an individual is HR-HPV positive. In accord with previous cost-effectiveness studies of screening in England, the success rate of treatment is assumed to be 95% for clearance of lesions, however, 16% of treated women are assumed to remain HPV positive[14].

Economic Assumptions

A cost-effectiveness analysis was conducted by comparing the incremental costs and outcomes over the lifetime of cohorts beginning screening in 2014 under the primary cytology and primary HPV protocols. Guidelines for the reference case of the National Institute for Health and Care Excellence (NICE) were followed. Costs were estimated from the perspective of the health care provider. Outcomes were

measured in terms of number of additional health care costs, cancers prevented, life years saved and quality-adjusted life years (QALYs) saved. A discount rate of 3.5% was used throughout. Costs were inflated to 2013/14 using the Hospital and Community Health prices index. Probabilistic sensitivity analysis was conducted incorporating uncertainty in both epidemiological and economic parameters.

Screening costs were obtained from previous economic analyses in which original data was collected at cervical screening sites in England[15–19]. Costs were inflated to 2013/2014 values using the hospital and community health index. Current cytology costs were also obtained for a sample laboratory taking part in the primary HPV pilot study. Costs were broken down according to initial sample collection; equipment and consumables; sample preparation and reading time; and other laboratory overhead costs. Historical economic were used to calculate an expected value for each cytology cost. In studies where overheads and other laboratory administrative costs were not reported, missing values were replaced with an average from studies in which costs were available. Given the improved economies in HPV testing technology over time and the change in costs that accompanies a switch to the primary HPV protocol, compared to HPV in triage, we do not inflate costs from all historical studies. Instead, we incorporate costs from recent studies from 2010 onwards[14,16,19], recommended costs of HPV tests according to the NHS supply chain in 2014 and costs reported by a site taking part in the primary HPV testing pilot. Where overheads and staff costs are missing for HPV testing, we augment costs using additional costs reported by the primary pilot site. A breakdown of all screening costs and sources can be found in Table 2 (described in detail Appendix A5).

The cost of cancer treatment was derived using the observed treatment preferences as a function of cervical cancer stage at diagnosis, as reported in the cervical cancer audit[20]:cone biopsy or loop excision, trachelectomy, hysterectomy alone, radiotherapy (with or without hysterectomy); chemotherapy (with or without hysterectomy); chemo-radiotherapy (with or without hysterectomy) (Table 2; described in detail in Appendix A5).

QALY weights for screening outcomes were based on previous values used in England[19,21–23], and more recent studies exploring QALY loss relating specifically to HPV primary screening in the Netherlands[24] and Australia[25]. We calculated a score for each combination of screening outcomes; the mean value is taken to be the mean score generated using utility scores from multiple sources. The 95% confidence intervals reflect the extreme utility scores generated in previous studies (Table 3). We implemented a quality of life detriment for 18 months following treatment for cervical cancer using the same approach as above and values taken from the literature[5,22,24,26–28]. Cancer mortality rates were calculated using the 1 and 5 years survival rates published by ONS; the data were used to parameterise an age-dependent mortality hazard function following diagnosis of cancer (described in appendix A3). In the model, women who survive beyond five years were assumed to avoid cervical cancer-related mortality, but incur a lifelong post cancer treatment quality of life detriment.

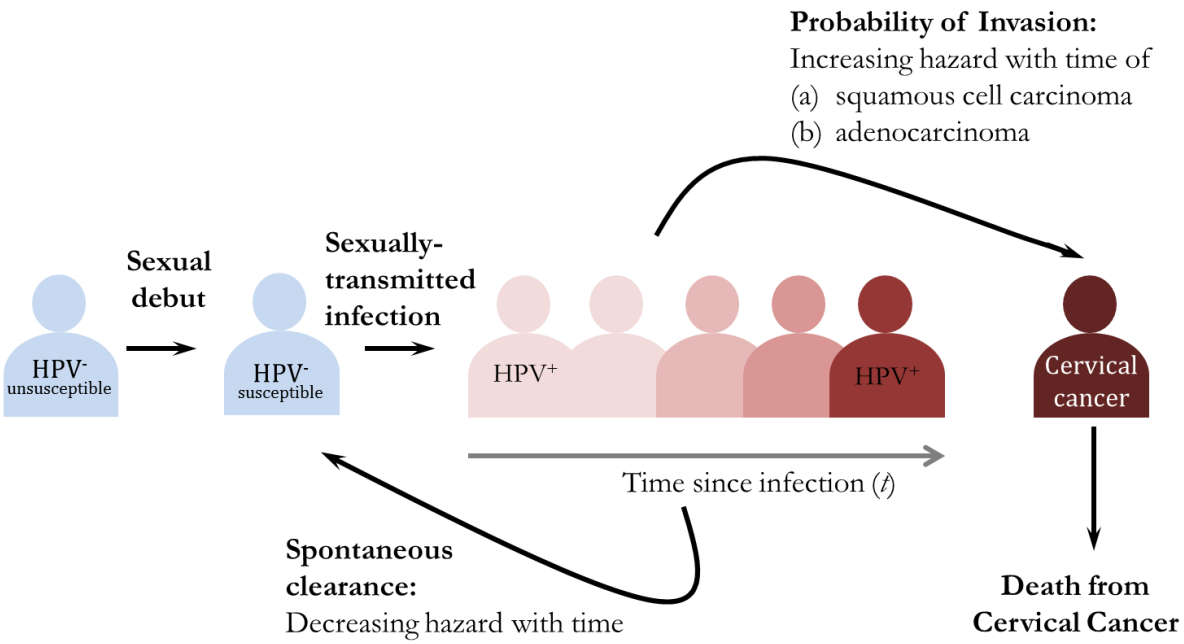


Figure 1: Model outline of HPV transmission and progression to cancer. The model simultaneously considers transmission of HPV-16, 18, 31, 33, 45, 51, 52 and 58.

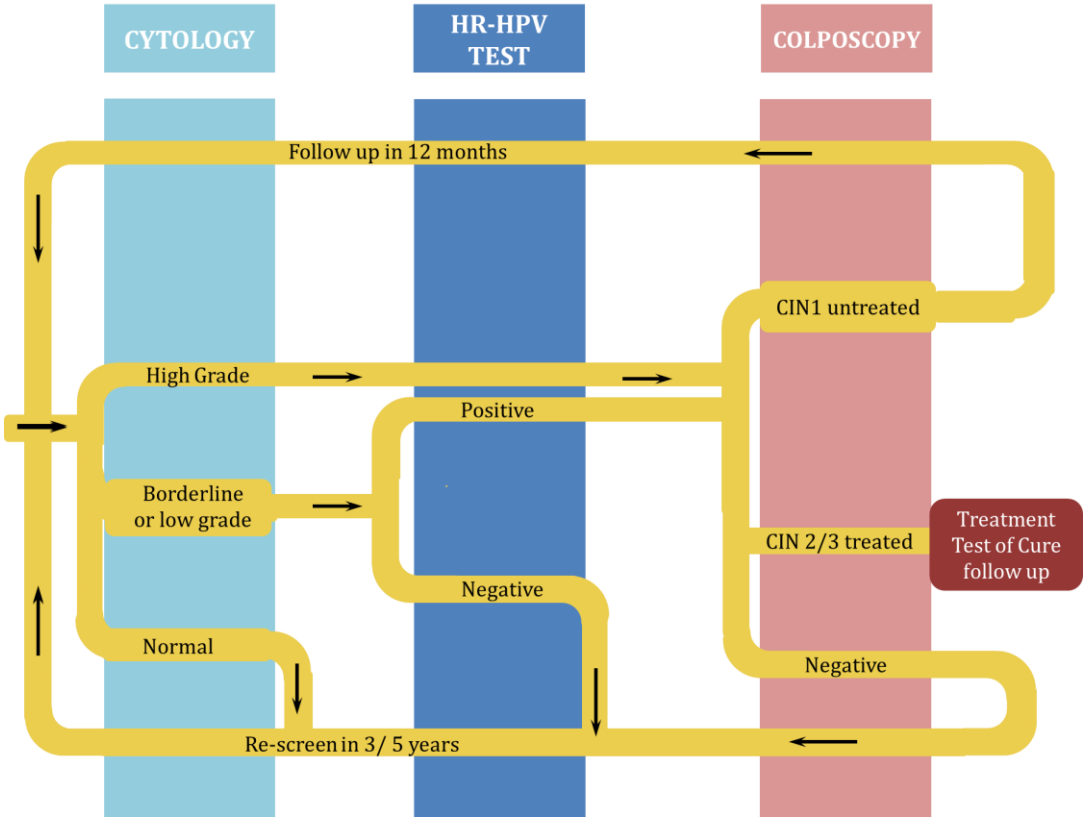


Figure 2: Primary cytology protocol – current screening practise.

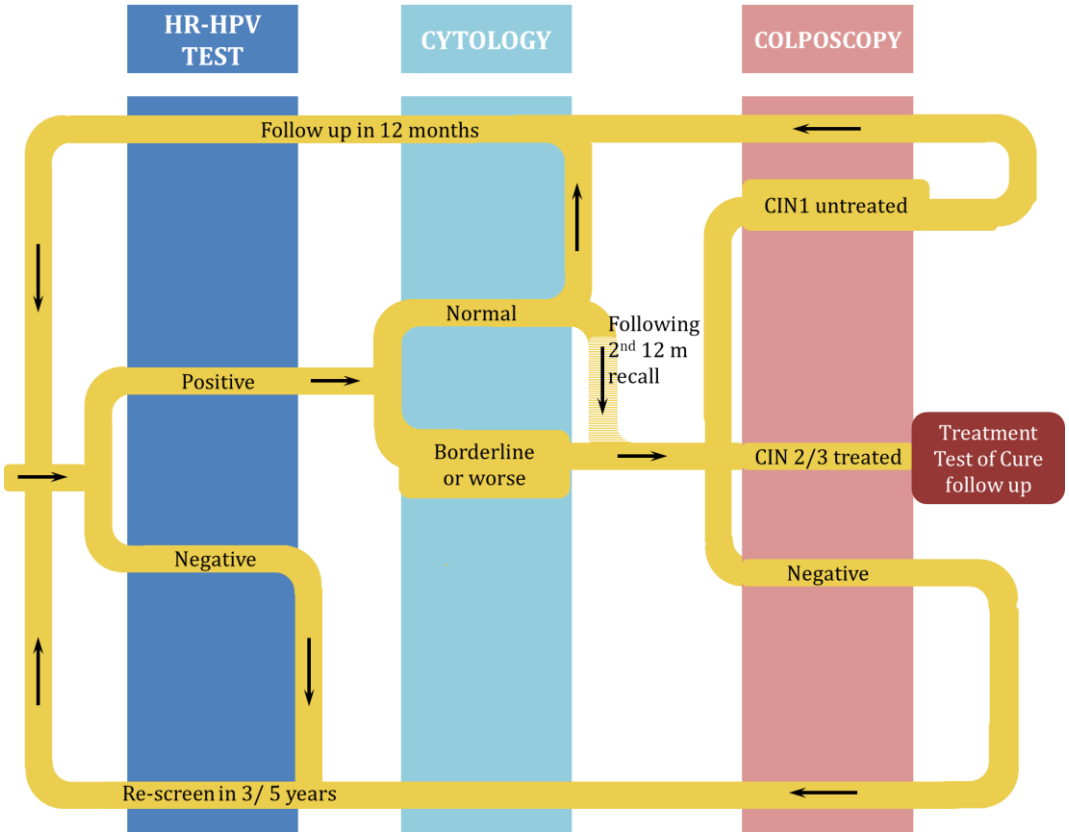


Figure 3: Primary HPV protocol

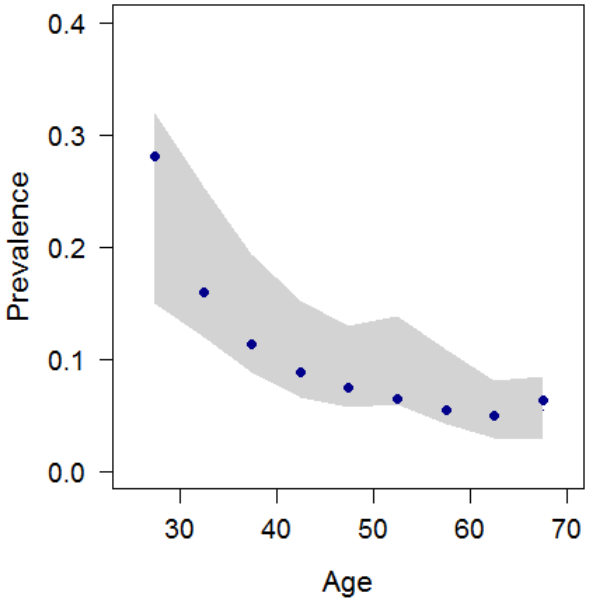


Figure 4: **Validation of model parameterisation**. HPV positivity projected by the model for women aged 25-70 undergoing primary HPV testing (grey shaded area = 95% interval) is parameterised using surveillance data collected by PHE that measures type-specific prevalence of HPV in women[9,10]. HPV positivity observed in the preliminary data from the primary HPV testing pilot sites (October 2014).

	Percentage attendance	Probability of normal outcome	Probability CIN1 detected	Probability CIN2 or worse detected
Current screening practise				
Borderline or Mild referral (n=21,977)	75.2%	55.4%	26.9%	17.7%
Moderate or worse referral (n=38,570)	78.0%	7.4%	8.1%	84.5%
Preliminary Primary HPV pilot outcomes				
Borderline or Mild referral (n=1473)	79.6%	66.1%	17.6%	16.4%
Moderate or worse referral (n=853)	88.0%	10.9%	6.1%	83.0%

Table1: Colposcopy outcomes under primary cytology algorithm (annual screening report 2012-2013), and preliminary outcomes form primary HPV pilot sites (October 2014).

Parameter	Costs	95 % range	Source
Screening			
Sample collection	15.31	(12.5, 18.63)	Karnon (2003), Moss (04), Kitchner (2011), LeGood (2012); Kitchner (2014)
HPV test per sample (includes consumables, equip,ment, staff time & other overheads)	9.75	(7.23, 13.03)	LeGood (2012); Kitchner (2014); NHS supplier chain (2014); Primary HPV pilot site (2014)
Cytology test per slide (includes consumables, equip,ment, staff time, other overheads)	18.15	(14.95, 22.02)	Karnon (2003), Moss (2004), Kitchner (2011), LeGood (2012); Kitchner (2014)
Treatment of pre cancer and cancers			
Colposcopy	151.18	(124.18, 184.08)	Martin-hirsch (2007)
Biopsy	79.84	(65.35, 97.71)	Sherlaw-Johnson (2004)
Excision	382.6	(313.89, 468.41)	Martin-hirsch (2007)
Hysterectomy	2583.5	(2222.28, 3039.77)	Martin-hirsch (2007)
Chemotherapy	5089	(4203.03, 6188.00)	Salter (2014)
Trachalectomy	5485.67	(4500.32, 6646.50)	Salter (2014)
Radiography	19078	(15709.73, 23126.39)	Salter(2014)
Stage 1	4,619		Salter (14); Cervical Cancer Audit (10)
Stage2	20,704	(17927.10, 23509.72)	Salter (14); Cervical Cancer Audit (10)
Stage 3	20,387	(17638.43, 23509.18)	Salter (14); Cervical Cancer Audit (10)
Stage 4	17,320	(14953.77, 20008.25)	Salter (14); Cervical Cancer Audit (10)

Table 2: Model inputs: economic parameters and sources.

	Utility loss per episode	95 % range	Sources
Screening outcomes			Simonella (2014);
Routine screen			Gold (1998) as used by Mandelblatt (2002) and de Kok (2014);
Negative cytology; Negative HPV	0.0001	(0.00002, 0.00023)	Myers (2007) as used in Elbasha (2007) and Kitchner (2014);
Abnormal result with routine recall			Insigna (2007);
Low grade cytology & negative HPV;	0.0011	(0.00023, 0.002)	TOMBOLA (2007)
Abnormal result with 12 month follow up			
Positive HPV & normal cytology	0.0040	(0.00023, 0.0089)	
Normal outcome at colposcopy			
Low grade cytology, positive HPV & normal colposcopy;			
High grade cytology & normal colp;			
Positive HPV, abnormal cytology & normal colposcopy	0.0147	(0.0015, 0.04)	
CIN1 outcome at colposcopy			
Low grade cytology, positive HPV & CIN1;			
High grade cytology & CIN1;			
Positive HPV, abnormal cytology & CIN1	0.0618	(0.005, 0.11)	
CIN2 outcome at colposcopy			
Low grade cytology, positive HPV & CIN2 or worse;			
High grade cytology & CIN2 or worse;			
Positive HPV, abnormal cytology & CIN2 or worse	0.0783	(0.003, 0.13)	
Cancer			Gold (1998), Stratton(2000) and Wolfson(1996)
stage 1	0.295	(0.19, 0.51)	as used in Goldie (2004), Kahn(2008),
stage 2	0.385	(0.33, 0.58)	deKok (2014) and Kitchner (14);
stage 3	0.440	(0.44, 0.58)	Myers (2004) as used by Elbasha (2007) and Jit (2011);
stage 4	0.520	(0.4, 0.64)	Klee (2000) and Korfage (2009)
post treatment			
stage 1	0.030	(0.01, 0.27)	
stage 2	0.065	(0.02, 0.32)	
stage 3	0.065	(0.02, 0.32)	
stage 4	0.205	(0.031, 0.53)	

Table 3: Model inputs: utility loss due to screening

Results

Clinical outcomes

A summary of clinical outcomes under the primary HPV and cytology protocols is shown in Table 4. The annual number of primary screening tests carried out is expected to increase by 4% under the standard primary HPV protocol from 3.03 million to 3.16 million per annum (Table 4); the largest increase is expected in women aged 25 to 35 and represents additional follow up testing for women found to be hpv positive but cytology negative (Figure 5).

Inevitably, the primary HPV protocol resulted in a large reduction in the absolute number of women undergoing cytological testing, from 2.999 to 0.305 million tests annually. One knock-on effect of this was that the proportion of women with non-negative cytology outcomes, among those undergoing cytology, increased from 10% under primary cytology protocol to 47% under primary HPV protocol (Figure 6). A more detailed breakdown of number of tests and outcomes is shown for each screening strategy in Tables 5 and 6.

The model predicted an 18% increase in the number of women attending colposcopy. There was a ~29% increase in the number of cases of CIN 2 or worse identified annually; reflecting ~18,000 additional cases detected per year through the screening programme (Tables 6 and 7; Figure 7). Over half of these additional incidences of CIN 2 or worse were identified in women under the age of 35 years. The model predicts an increase in the 'efficiency of screening' as measured by number of women screened to identify a single case of CIN2 or worse; 50 women need to be screened using the primary cytology protocol, compared to 40 under the primary HPV protocol, to identify a single case if CIN2 or worse.

The rare nature of cervical cancer means that the best fitting model simulations cover a wide range of scenarios for cancer incidence when we combine cases of squamous cell carcinoma and adenocarcinoma for HPV types 16,18, 31, 33, 51, 52 and 58 (Figure 8), however, the model predicts a median decrease in the incidence of cervical cancer of 310 cases per year (IQR(-647, 1379)) (Table 5). Despite the noisy model projections for both scenarios, we see a consistent decrease in cancer incidence within each 5 year age-band for women of screening age; with the largest benefits expected in women from aged 30 years onwards (Figure 8). In terms of the 'efficiency of primary screening', we find that primary HPV protocol requires an additional 397 primary screens per cancer case avoided. This reduction in cancer incidence leads to a median the saving of 73 lives per year (IQR(-168, 348)).

Economic outcomes

The primary HPV protocol is expected to have lower net costs compared to the primary cytology protocol (Table 5); the benefit of avoiding cytological screens, which are more expensive than HPV tests, outweigh the cost of increased primary screens, colposcopies and treatments. The annual screening costs are predicted to be £134 million under a primary cytology strategy and £120 million under a primary HPV strategy. In terms of total health-related costs, including the cost of cancer treatment, this increases to £153 million under primary cytology, and £136 million under primary HPV; resulting in a median saving of £15.8 million (IQR=(2.7m, 27m)). The median discounted cost savings over the lifetime is forecast to be £14 per woman.

The primary HPV is expected to be life-saving, the median saving of 73 lives; with cervical cancer resulting in 520 and 461 deaths per year under primary cytology and primary HPV, respectively. This life saving translates into a median discounted per-woman life year saving of 0.0018 (-0.0043, 0.0082).

The model predicts that a switch to primary HPV protocol would lead to a median increase in the discounted number of quality adjusted life years (QALYs) lost, per women, of 0.0026 (interquartile range= (-0.0013, 0.0064). The gain in life years and reduction in cancers is counteracted by the countered by a larger increase in the utility cost of increased primary testing, colposcopies and women being treated for CIN2 or worse under primary HPV testing. The large confidence intervals predominantly reflect the variation associated with screening-related QALY detriments, in addition to the model uncertainty surrounding the projected number of cancers. In our primary analysis, we use QALY weights that are an average of those reported in the literature and assume a normal distribution to cover all reported values, however, this potentially unfairly skews the qalys towards higher values. We find that some older studies[22,23] report a qaly detriment associated with colposcopy that is of the order of 16-37 fold higher than that of more recent studies[24,25]. To explore the sensitivity of our results to screening-related QALY detriments, we repeat the analysis using QALY values from the study reporting the strongest ("Insigna Basis"[23]) and weakest ("Simonella Basis" [25]) screening-related detriments. Using the Simonella QALY basis, we find a small median gain in discounted per-woman lifetime QALYs of 0.0005 associated with a move from the primary cytology to primary HPV protocol, while, using the Insinga study, gives a median loss of 0.0033.

Extended screening interval

Evidence for the stronger negative predictive power of HPV over cytology and the concern regarding the over-testing in young women, in whom there is a high prevalence of HPV infection, have led to a discussion in the health care community regarding the extension of the standard screening recall interval associated with primary HPV testing. We consider the impact of increasing the recall interval, following a negative primary HPV screen, to 5 years for all women regardless of age (5 year primary HPV protocol). This fixed interval compares to current practise whereby women under 50 years are recalled at 3 year intervals, and women over 50 are recalled at 5 year intervals.

As we might expect, the model predicts a 17% decrease in the number of primary tests carried out when the recall interval is extended from 3 to 5 years for women under 50 (from 3.034 to 2.514 million tests per year). The number of colposcopies is predicted to remain unchanged with a move from primary cytology to primary HPV with 5 year recall; however, the model predicts an increase in the number of CIN2 or worse cases detected from 61,504 to 70,400 per year. The increased 'rate' of detection per colposcopy under a 5 year protocol arises from the increased proportion of women attending colposcopy following a moderate or severe cytological referral. Overall, the increased detection and subsequent treatment of precancerous lesions results in a drop in cancer incidence of 159 cases per year under the 5 year primary HPV protocol, saves 54 lives per year, and leads to a discounted per-woman lifeyear saving of 0.0008.

Moving from primary cytology to primary HPV testing, in combination with a regular 5 year screening interval, would lead to a substantial total health-care cost saving of £35 million (22.4m, 47.2m). The

annual screening costs are expected to be £97.7 million. The discounted lifetime cost saving per women is estimated to be £38 (25,49).

In line with the observations for standard primary HPV protocol, when we use our mixed QALY weighting basis, the gain in life years associated with a switch to a 5 year primary HPV protocol are dominated by the QALY detriment resulting from increased detection and treatment of CIN2 cases; the modelling predicts a median discounted per-woman lifetime QALY loss of 0.001 (-0.0047, 0.0028). As before, we show that the resulting QALY outcome is highly sensitive to the screening-associated QALY weights used. The Simonella basis for screening-related QALYs leads to a median gain in discounted per-woman lifetime QALYs of 0.0052, while the Insinga basis leads to a median loss in discounted per-woman lifetime QALYs of 0.0009.

Summary

The modelling work presented here predicts that a move from the current primary cytology to a primary HPV screening protocol will be both life-saving and cost-saving. However, the benefits as measured by quality adjusted life years are more difficult to determine due to the uncertainty associated with screening associated quality of life detriments. We find that a switch to primary HPV screening can be shown to result in: (i) QALY gains when using screening-associated quality of life detriments measured in a recent study looking explicitly at primary HPV testing by Simonella and colleagues[25]; but also (ii) QALY losses when using more severe quality of life detriments screening-associated as reported by Insinga and colleagues[23],where life year gains are obscured by QALY detriments resulting from significant increases in colposcopy referrals and identification and treatment of precancerous lesions.

In terms of clinical outcomes, moving from the current cervical screening protocol to one employing primary HPV testing is expected to: (i) increase the number of primary screening tests carried out; (ii) increase the number of women referred to colposcopy; and (iii) increase the number of lesions of grade 2 or worse identified and treated through colposcopy. The model projects a positive impact on cervical cancer incidence and cancer-related mortality.

The impact of increasing the standard recall interval, following a negative primary HPV screen, to 5 years for all women, regardless of age, is also considered within the primary HPV protocol. The switch from a primary cytology to 5 year primary HPV protocol is expected to: (i) reduce cancer incidence; (ii) reduce cancer-related deaths; and (iii) reduce costs. As above, the predicted change in QALYs is a mixed bag; the optimistic Simonella basis predicts a QALY gain, while the more severe Insinga basis predicts a QALY loss.

The model predicts a sizable total health-care cost saving of £35 million (22.4m, 47.2m) with a switch from the current practise primary cytology protocol to the 5 year primary HPV protocol, compared to a saving of £15.8 million (2.7m, 27m) associated with a switch to the standard primary HPV protocol. The median reduction in cervical-cancer related deaths is predicted to be 54 and 73, respectively, following a switch to the 5 year- and standard-, primary HPV protocols. Despite the smaller life-years saving, the trade-off between screening- and cancer- related QALY losses means that a switch to a 5 year primary HPV protocol is more favourable in terms of net quality-adjusted life years than a switch to standard primary HPV testing. The median QALY loss predicted, using an averaged QALY weighting basis, for a switch from current practise to a 5 year recall primary HPV protocol is 0.0010, compared to a QALY loss of 0.0026 associated with a switch from current practise to the standard HPV protocol.

Model Limitations

The model explicitly considers HPV strains 16, 18, 31, 33, 45, 51, 52 and 58, representing the most prevalent strains that are associated with cervical cancer in England. However, commercially available test, such as the commonly used HC2 assay, will also detect cases of hpv-35, 39, 56, 59 and 68. There are also reports that HPV testing may react to non HR-HPV test, however, the validation of model outcomes against preliminary data from the HPV primary pilot give us confidence that we do not underestimate HR-HPV positivity.

Model projections give a large uncertainty range around cancer incidence. This uncertainty is in part explained by the additive uncertainty arising from combining 16 distinct cancer-causing processes– eight hpv strains leading to either squamous cell carcinoma or adenocarcinomas. The rare nature of non hpv 16/18-related cancers means that the underlying parameters can be difficult to constrain for hpv strains other than 16 and 18. Conservatively, the model simulations cover a broad range of scenarios for each HPV type.

In this work, we use the economic costs taken from historical economic analyses of screening in England, and inflate to 2014 values. The limitations of inflating historical costs are that we don't necessarily capture the reduction in technology costs over time. Economies of scale also suggest that a switch to primary screening is likely to result in a reduction in the per sample cost of a hpv test. Overall, this is expected to lead to a further cost saving associated with a switch to primary HPV testing. A more detailed study of work flow and costs in the context of primary HPV testing is planned by the primary HPV pilot screening committee that will provide further insight into the expected changes.

The utility detriment associated with cervical screening is not well defined, reflected in the diverse estimates for QALY loss weights reported in the literature. This is particularly true for primary HPV related screening. In this work, we use a sensitivity approach that captures the extreme values reported in the literature to show that the choice of published screening-associated QALY loss values can determine whether an intervention is beneficial or detrimental. The work highlights a need for further study of QALY loss associated with screening, in order to appropriately judge the increase in colposcopy and treatment of precancerous CIN2 lesions we are willing to accept in order to reduce the incidence and death related to cervical cancer.

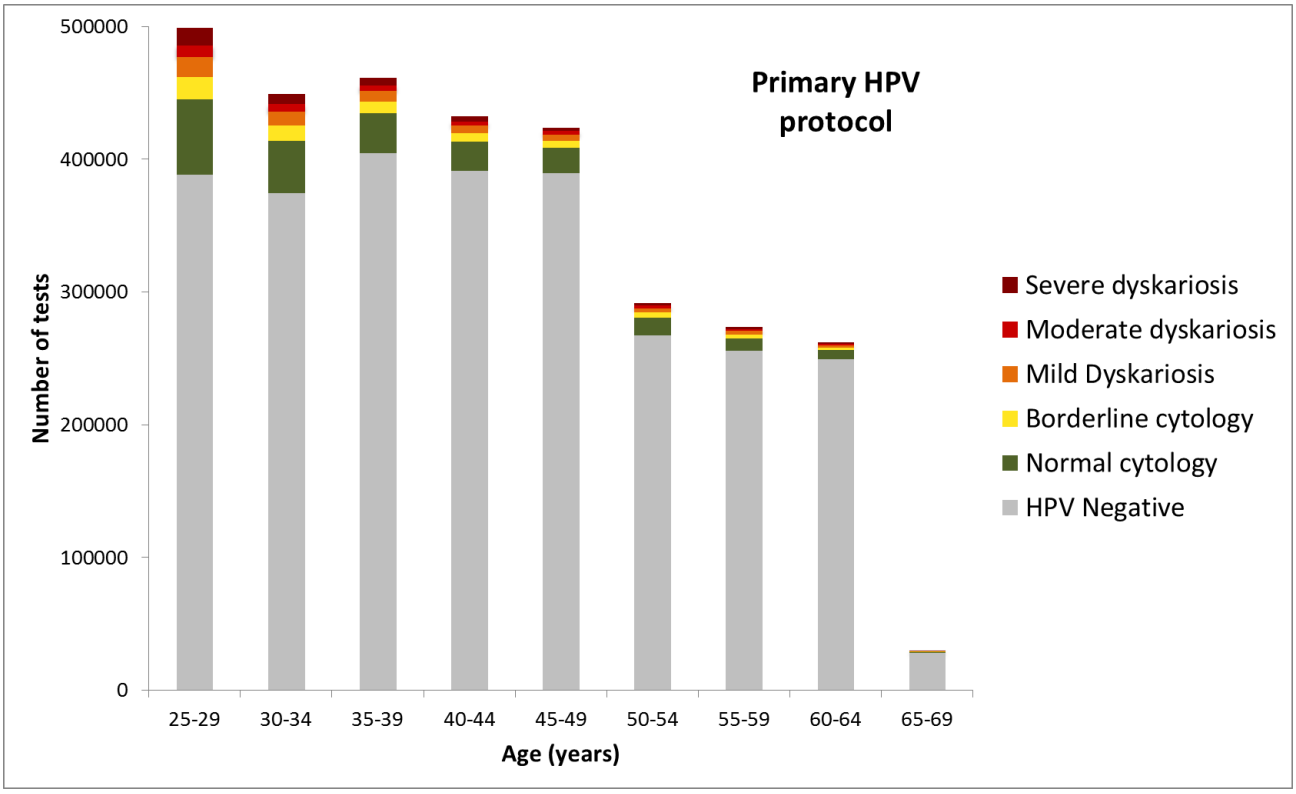
The current analysis is based on a static model of infection, this means that we are unable to incorporate changes in male prevalence that might arise following vaccination due to herd immunity; and limits the projections that we can make about the suitability of HPV testing to an unvaccinated female population. The introduction of a national HPV vaccination programme, in 2008, means that it is relevant to consider the implications of vaccination on HPV prevalence and disease incidence as vaccinated cohorts approach screening age. The model is currently being developed to include a dynamic disease transmission element that will allow further work to consider the optimal screening protocol for vaccinated women. It is possible that alternative scenarios, in which the screening interval is further extended or HPV 16/ 18 genotyping is included, might prove more cost-effective than primary HPV testing alone.

	Primary cytology protocol	**Primary HPV protocol** with 3 year recall for women under 50, and 5 year recall otherwise
Number of cytology tests	**3,034,422**	**372,980**
	(2958719, 3138301)	(276902, 547579)
Normal cytology	**2,713,165**	**196,037**
	(2615090, 2786752)	(118158, 324281)
Borderline changes	**136,302**	**57,997**
	(99507, 176631)	(39417, 92495)
Mild dyskariosis	**94,194**	**52,260**
	(60763, 134127)	(27734, 88093)
Moderate dyskariosis	**40,029**	**28,673**
	(11360, 92688)	(6126, 80683)
Severe dyskariosis	**50,657**	**37,938**
	(16230, 112887)	(11304, 89923)
Number of HPV tests	**282,571**	**3,157,452**
	(203349, 391713)	(3079357, 3273993)
HPV negative	**180,767**	**2,748,233**
	(128577, 255488)	(2641853, 2818555)
HPV positive	**101,785**	**409,200**
	(56110, 176393)	(298156, 605681)
Number of colposcopies	**147,925**	**174,996**
	(68894, 317668)	(106505, 301443)
Normal	**61,665**	**63,398**
	(36021, 101136)	(42764, 100978)
CIN 1	**25,169**	**32,686**
	(12373, 50388)	(21843, 51877)
CIN 2 or worse	**61,054**	**78,875**
	(18710, 166379)	(34027, 167638)

Table 4: Summary of clinical outcomes and resource usage (mean and 95%CI). Number of tests calculated assuming an age distribution as observed by ONS in 2013.

	Primary cytology protocol	**Primary HPV protocol** with 3 year recall for women under 50, and 5 year recall otherwise	**Saving under primary HPV**	**Primary HPV protocol** with 5 year recall for all women	**Saving under primary HPV with 5 year recall**
Annual screening-associated costs (£000)	**134,173** (122855, 145382)	**120,479** (112413, 130635)	**13,078** (2924, 22814)	**97,726** (91366, 106906)	**33,958** (23749, 44166)
Annual total health costs (£000) (including cost of cervical cancers)	**153,391** (139306, 164510)	**136,707** (126156, 147393)	**15,756** (2716, 27990)	**114,196** (104471, 126831)	**35,711** (22381, 47182)
Discounted lifetime cost per women (£) (including cost of cervical cancers)	**160** (146, 172)	**145** (134, 157)	**14** (-1, 27)	**121** (108, 131)	**38** (25, 49)
Annual incidence of cervical cancer	**2123** (1208, 3290)	**1828** (1016, 2738)	**310** (-647, 1379)	**1999** (1152, 3022)	**159** (-820, 1070)
Deaths related to cervical cancer (/year)	**520** (290, 812)	**461** (235, 700)	**73** (-168, 348)	**475** (276, 732)	**54** (-192, 272)
Discounted life years lost to cervical cancer per women	**0.0157** (0.0092, 0.0239)	**0.0146** (0.0079, 0.0212)	**0.0018** (-0.0043, 0.0082)	**0.0153** (0.0085, 0.0224)	**0.0008** (-0.0063, 0.0076)
Discounted quality-adjusted life years lost due to cancer and screening	**0.0136** (0.0105, 0.0165)	**0.0160** (0.0128, 0.0198)	**-0.0026** (-0.0064, 0.0013)	**0.0144** (0.0113, 0.0179)	**-0.0010** (-0.0047, 0.0028)
Discounted quality-adjusted life years lost due to cancer and screening, using Simonella basis for screening-related QALY detriment	**0.0060** (0.0037, 0.0080)	**0.0055** (0.0033, 0.0073)	**0.0005** (-0.0013, 0.0026)	**0.0052** (0.0032, 0.0076)	**0.0004** (-0.0018, 0.0025)
Discounted quality-adjusted life years lost due to cancer and screening, using Insigna basis for screening-related QALY detriment	**0.0195** (0.0151, 0.0225)	**0.0225** (0.0184, 0.0262)	**-0.0033** (-0.0004, 0.0064)	**0.0199** (0.0.0162, 0.0237)	**-0.0009** (-0.0040, 0.0020)

Table 5: Summary of costs and health outcomes (mean and IQR).



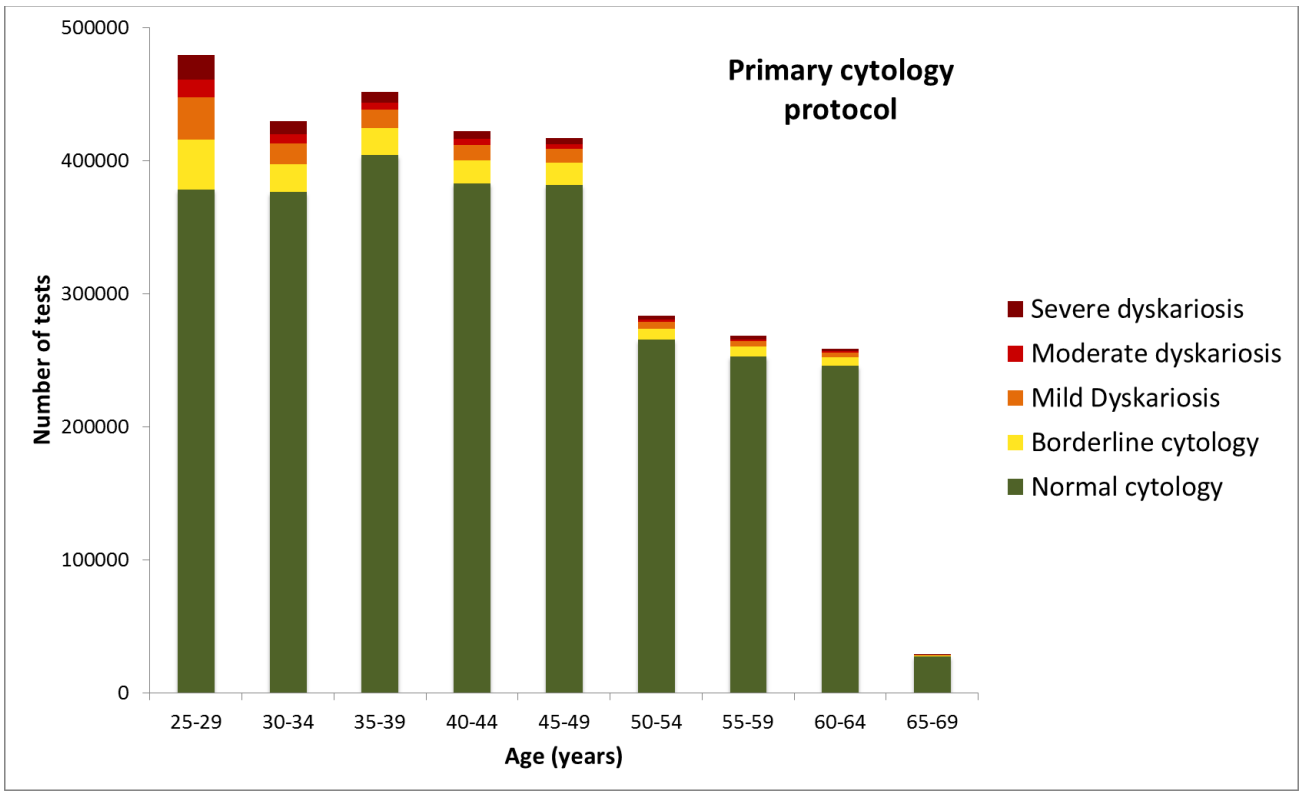


Figure 5: Number of women tested and predicted outcome under primary HPV protocol and primary cytology protocol, assuming age distribution in England as in 2013.

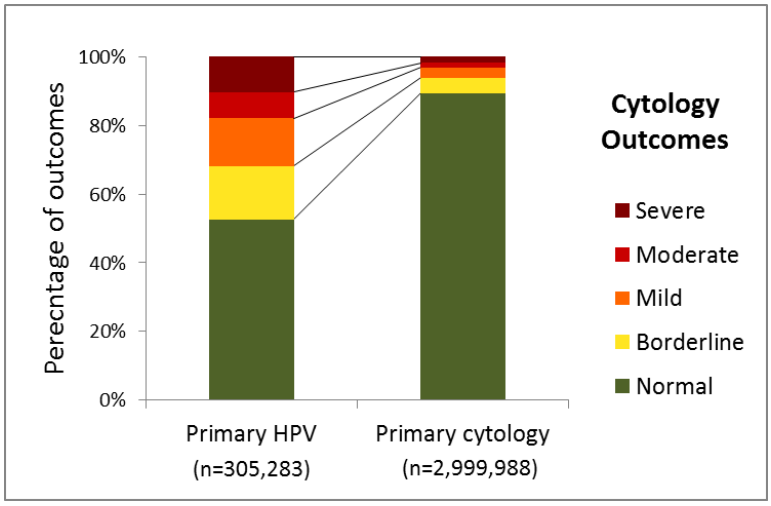


Figure 6: Outcome of cytology tests under primary cytology protocol (left) and primary hpv protocol.

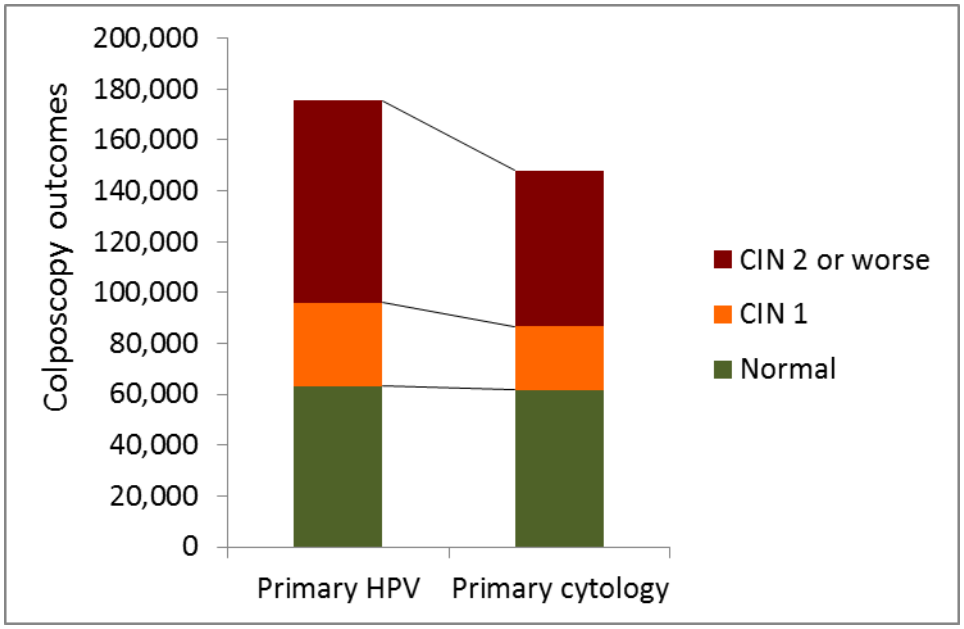


Figure 7: Predicted number and outcome of colposcopy tests undertaken.

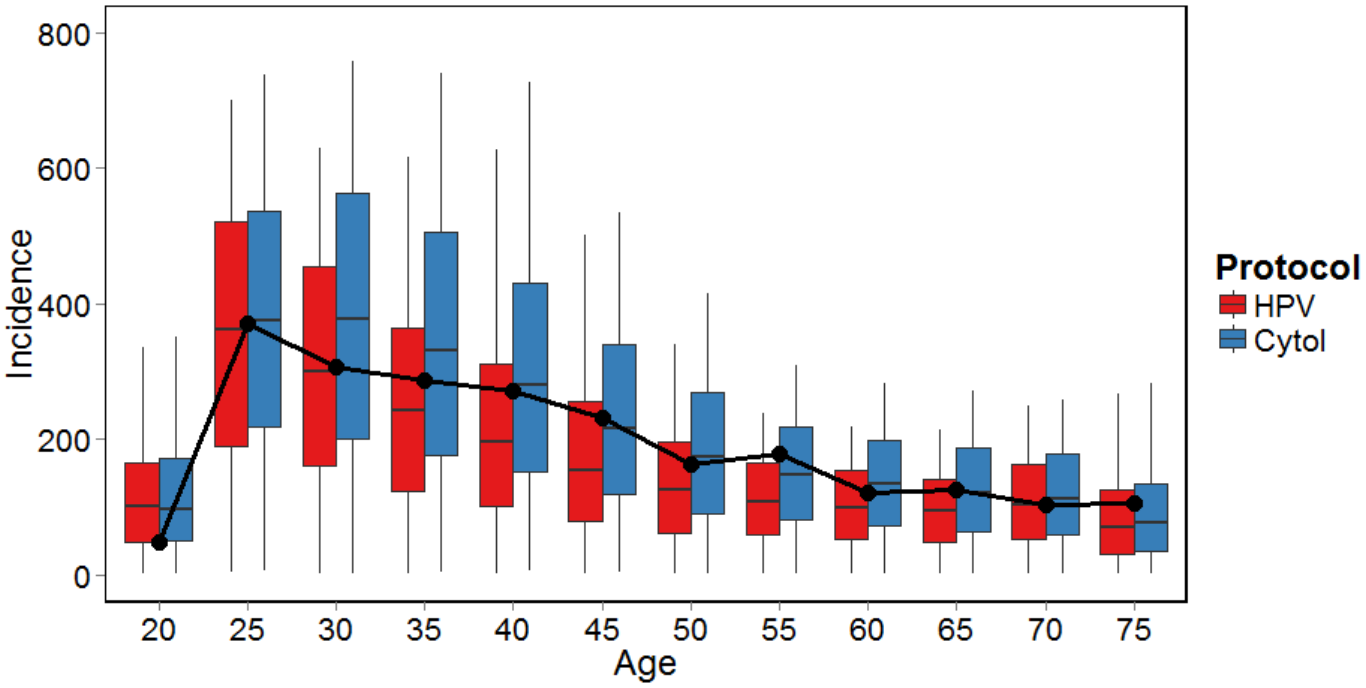


Figure 8: Cervical cancer incidence observed in 2012 (points) compared to model outcomes under primary cytology and primary HPV protocols. Boxes represent the interquartile range range of model predictions for cancer incidence (primary cytology =blue; primary hpv= red).

	Age									
	25 - 29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	**All**
Number of cytology tests	**110,338**	**74,439**	**56,436**	**40,700**	**34,161**	**24,617**	**18,132**	**12,754**	**1,404**	**372,980**
	(81673, 160273)	(55816, 107816)	(42196, 82586)	(30144, 60559)	(25664, 50617)	(18245, 36491)	(13102, 27486)	(9082, 19512)	(980, 2239)	(276902, 547579)
Normal cytology	56,149	38,853	30,049	21,769	18,788	13,401	9,535	6,754	738	196,037
	(34030, 91710)	(23523, 63728)	(18463, 49045)	(13011, 36063)	(11406, 31176)	(7878, 22697)	(5553, 16599)	(3877, 11944)	(418, 1320)	(118158, 324281)
Borderline changes	16,803	11,681	8,904	6,358	5,422	3,928	2,769	1,922	211	57,997
	(11484, 26620)	(7988, 18443)	(6127, 14083)	(4284, 10227)	(3694, 8679)	(2633, 6327)	(1824, 4561)	(1252, 3191)	(131, 364)	(39417, 92495)
Mild dyskariosis	15,371	10,609	7,970	5,733	4,689	3,426	2,519	1,748	193	52,260
	(8315, 25364)	(5605, 17747)	(4224, 13443)	(3074, 9736)	(2412, 8159)	(1760, 5946)	(1316, 4321)	(922, 3039)	(106, 341)	(27734, 88093)
Moderate dyskariosis	8,921	5,734	4,223	3,046	2,426	1,815	1,422	980	106	28,673
	(1923, 24227)	(1233, 16047)	(915, 11935)	(660, 8728)	(505, 6941)	(376, 5372)	(291, 4207)	(199, 2929)	(23, 297)	(6126, 80683)
Severe dyskariosis	13,085	7,552	5,281	3,785	2,826	2,038	1,877	1,341	153	37,938
	(3462, 32565)	(2337, 17191)	(1719, 11989)	(1246, 8603)	(927, 6476)	(630, 4863)	(543, 4605)	(393, 3272)	(47, 359)	(11304, 89923)
Number of HPV tests	504,798	457,864	467,404	436,767	427,406	294,728	275,439	263,416	29,630	3,157,452
	(483460, 535001)	(442663, 481441)	(457877, 482145)	(428105, 450261)	(420751, 437499)	(288540, 304578)	(270302, 282813)	(258880, 269371)	(28778, 30883)	(3079357, 3273993)
HPV negative	388,228	374,410	404,511	391,256	389,692	266,977	255,653	249,489	28,018	2,748,233
	(355827, 409882)	(355911, 386154)	(387233, 415050)	(380439, 398100)	(379801, 395972)	(260879, 271345)	(249820, 259920)	(244415, 253595)	(27527, 28538)	(2641853, 2818555)
HPV positive	116,568	83,452	62,891	45,508	37,712	27,748	19,784	13,925	1,611	409,200
	(85431, 170019)	(61117, 121951)	(46173, 92900)	(32821, 68080)	(27781, 56038)	(20052, 41858)	(14016, 30503)	(9701, 21718)	(1066, 2614)	(298156, 605681)
Number of colposcopies	49,804	36,715	26,967	19,441	15,434	11,690	8,385	5,850	711	174,996
	(27944, 89477)	(23139, 61167)	(17041, 44960)	(12287, 32599)	(9750, 26343)	(7411, 19910)	(5030, 15266)	(3476, 10469)	(427, 1253)	(106505, 301443)
Normal	16,874	13,543	10,067	7,260	5,863	4,483	2,980	2,070	259	63,398
	(11283, 26449)	(9258, 21420)	(6870, 16105)	(4904, 11657)	(4017, 9301)	(2995, 7218)	(1950, 4916)	(1328, 3456)	(158, 456)	(42764, 100978)
CIN 1	8,800	6,959	5,172	3,722	2,998	2,290	1,542	1,069	133	32,686
	(5745, 13918)	(4733, 10899)	(3537, 8125)	(2509, 5962)	(2035, 4772)	(1521, 3665)	(1000, 2531)	(681, 1776)	(80, 229)	(21843, 51877)
CIN 2 or worse	24,126	16,208	11,724	8,454	6,569	4,911	3,858	2,707	319	78,875
	(9264, 53662)	(7279, 33202)	(5378, 24070)	(3899, 17316)	(3006, 13848)	(2288, 10361)	(1627, 8565)	(1137, 5950)	(148, 665)	(34027, 167638)

Table 6: Model-generated number and outcome of HPV, cytology and colposcopy tests, per annum, under a primary HPV protocol; mean (lower & upper bound). Resident female population size and age demographic as observed in England in 2013 (source: ONS).

	Age									
	25 - 29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	**All**
Number of cytology tests	**478,489**	**429,077**	**451,045**	**421,524**	**416,611**	**283,045**	**267,892**	**258,378**	**28,361**	**3,034,422**
	(457843, 507404)	(414645, 450578)	(441944, 463634)	(413615, 432915)	(409993, 424870)	(277068, 290522)	(262543, 274852)	(253334, 264273)	(27734, 29253)	(2958719, 3138301)
Normal cytology	378,522	376,087	403,957	382,686	381,388	265,191	252,657	245,673	27,004	2,713,165
	(348961, 397886)	(360253, 387322)	(388912, 414118)	(372955, 390835)	(371664, 389701)	(259319, 270441)	(246551, 257897)	(240118, 250872)	(26357, 27680)	(2615090, 2786752)
Borderline changes	37,488	21,266	20,361	17,561	16,755	8,409	7,281	6,482	700	136,302
	(29423, 45705)	(15891, 27523)	(14760, 26489)	(12555, 23095)	(11753, 22042)	(5649, 11526)	(4802, 10112)	(4224, 9150)	(450, 989)	(99507, 176631)
Mild dyskariosis	31,609	15,158	13,680	11,312	10,331	4,715	3,879	3,174	336	94,194
	(22313, 41186)	(9520, 22862)	(8697, 19805)	(7227, 16336)	(6412, 14870)	(2637, 7424)	(2121, 6093)	(1665, 5021)	(170, 529)	(60763, 134127)
Moderate dyskariosis	13,549	7,123	5,804	4,491	3,803	2,137	1,718	1,274	132	40,029
	(5406, 28865)	(1604, 17734)	(1442, 13747)	(1081, 10287)	(913, 8238)	(373, 5624)	(306, 4532)	(215, 3315)	(18, 345)	(11360, 92688)
Severe dyskariosis	17,312	9,434	7,234	5,465	4,324	2,584	2,349	1,767	188	50,657
	(5451, 39302)	(2775, 21335)	(2510, 15549)	(1885, 11645)	(1483, 8970)	(819, 5941)	(733, 5647)	(517, 4070)	(57, 428)	(16230, 112887)
Number of HPV tests	**81,403**	**47,987**	**42,580**	**35,145**	**31,764**	**16,646**	**13,985**	**11,800**	**1,261**	**282,571**
	(61055, 107277)	(32981, 71057)	(30713, 59336)	(25687, 48056)	(23420, 42622)	(11112, 24721)	(9539, 20212)	(7999, 16645)	(845, 1788)	(203349, 391713)
HPV negative	52,065	27,352	26,422	23,775	22,306	10,136	9,252	8,527	932	**180,767**
	(40544, 68381)	(18205, 43000)	(18415, 37654)	(16985, 32927)	(16053, 29963)	(6429, 15848)	(5858, 13918)	(5473, 12434)	(614, 1363)	(128577, 255488)
HPV positive	29,336	20,632	16,156	11,368	9,455	6,508	4,730	3,271	329	**101,785**
	(16062, 49235)	(11312, 35428)	(9105, 27481)	(6415, 20079)	(5203, 17175)	(3519, 12094)	(2554, 8515)	(1765, 5768)	(175, 618)	(56110, 176393)
Number of colposcopies	**46,262**	**28,682**	**22,445**	**16,370**	**13,468**	**8,580**	**6,761**	**4,856**	**500**	**147,925**
	(21996, 99207)	(12842, 62011)	(10765, 46499)	(7744, 34843)	(6328, 28615)	(3813, 19260)	(2996, 15302)	(2186, 10796)	(225, 1133)	(68894, 317668)
Normal	20,087	11,195	9,242	6,983	6,051	3,390	2,601	1,917	197	**61,665**
	(12722, 31512)	(6317, 18978)	(5345, 14856)	(3887, 11227)	(3341, 9867)	(1845, 6076)	(1437, 4604)	(1027, 3647)	(100, 369)	(36021, 101136)
CIN 1	7,373	5,104	3,951	2,787	2,285	1,583	1,181	822	83	**25,169**
	(3542, 14711)	(2501, 10169)	(2004, 7586)	(1399, 5536)	(1149, 4736)	(776, 3339)	(577, 2497)	(385, 1633)	(40, 181)	(12373, 50388)
CIN 2 or worse	18,797	12,378	9,248	6,595	5,127	3,603	2,975	2,112	218	**61,054**
	(5279, 52470)	(3784, 33045)	(3031, 23982)	(2167, 17688)	(1708, 14177)	(1148, 10125)	(906, 8415)	(620, 5856)	(65, 620)	(18710, 166379)

Table 7: Model-generated number and outcome of cytology and HPV tests, per annum, under a primary cytology protocol; mean (lower & upper bound). Resident female population size and age demographic as observed in England in 2013 (source: ONS).

References

1. Arbyn M, Sasieni P, Meijer CJLM, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: A summary of meta-analyses. *Vaccine*. 2006;24. doi:10.1016/j.vaccine.2006.05.117.

2. Cuzick J, Arbyn M, Sankaranarayanan R, et al. Overview of Human Papillomavirus-Based and Other Novel Options for Cervical Cancer Screening in Developed and Developing Countries. *Vaccine*. 2008;26. doi:10.1016/j.vaccine.2008.06.019.

3. Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ*. 2008;337:a1754. doi:10.1136/bmj.a1754.

4. Hoyer H, Scheungraber C, Kuehne-Heid R, et al. Cumulative 5-year diagnoses of CIN2, CIN3 or cervical cancer after concurrent high-risk HPV and cytology testing in a primary screening setting. *Int J Cancer*. 2005;116:136-143. doi:10.1002/ijc.20955.

5. Jit M, Chapman R, Hughes O, Choi YH. Comparing bivalent and quadrivalent human papillomavirus vaccines: economic evaluation based on transmission model. *Bmj*. 2011;343(sep27 1):d5775-d5775. doi:10.1136/bmj.d5775.

6. Canfell K, Barnabas R, Patnick J, Beral V. The predicted effect of changes in cervical screening practice in the UK: results from a modelling study. *Br J Cancer*. 2004;91(3):530-536. doi:10.1038/sj.bjc.6602002.

7. Sonnenberg P, Clifton S, Beddows S, et al. Prevalence, risk factors, and uptake of interventions for sexually transmitted infections in Britain: Findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). *Lancet*. 2013;382(13):1795-1806. doi:10.1016/S0140-6736(13)61947-9.

8. Kitchener HC, Almonte M, Wheeler P, et al. HPV testing in routine cervical screening: cross sectional data from the ARTISTIC trial. *Br J Cancer*. 2006;95(1):56-61. doi:10.1038/sj.bjc.6603210.

9. Howell-Jones R, de Silva N, Akpan M, et al. Prevalence of human papillomavirus (HPV) infections in sexually active adolescents and young women in England, prior to widespread HPV immunisation. *Vaccine*. 2012;30(26):3867-3875. doi:10.1016/j.vaccine.2012.04.006.

10. Howell-Jones R, Bailey a, Beddows S, et al. Multi-site study of HPV type-specific prevalence in women with cervical cancer, intraepithelial neoplasia and normal cytology, in England. *Br J Cancer*. 2010;103(2):209-216. doi:10.1038/sj.bjc.6605747.

11. Desai S, Chapman R, Jit M, et al. Prevalence of human papillomavirus antibodies in males and females in England. *Sex Transm Dis*. 2011;38(7):622-629. doi:10.1097/OLQ.0b013e31820bc880.

12. Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. *Br J Cancer*. 2003;89(1):88-93. doi:10.1038/sj.bjc.6600974.

13. Kelly RS, Patnick J, Kitchener HC, Moss SM. HPV testing as a triage for borderline or mild dyskaryosis on cervical cytology: results from the Sentinel Sites study. *Br J Cancer*. 2011;105(7):983-988. doi:10.1038/bjc.2011.326.

14. Legood R, Smith M, Lew J-B, et al. Cost effectiveness of human papillomavirus test of cure after treatment for cervical intraepithelial neoplasia in England: Economic analysis from NHS Sentinel Sites Study. *BMJ*. 2012;345(7882):e7086.

15. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. *Health Technol Assess (Rockv)*. 2004;8(20):1-90.

16. Kitchener HC, et al. MAVARIC - A comparison of automation-assisted and manual cervical screening: A randomised controlled trial. *Health Technol Assess (Rockv)*. 2011;15(3):1-176.

17. Moss SM, Gray A, Marteau T, Legood R, Henstock E, Maissi E. Evaluation of HPV / LBC Cervical Screening Pilot Studies Report to the Department of Health (Revised October 2004). 2004.

18. Martin-Hirsch P, Rash B, Martin A, Standaert B. Management of women with abnormal cervical cytology: Treatment patterns and associated costs in England and Wales. *BJOG An Int J Obstet Gynaecol*. 2007;114:408-415. doi:10.1111/j.1471-0528.2007.01261.x.

19. Kitchener HC, Canfell K, Gilham C, Sargent A, Roberts C, Desai M. The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England. *Health Technol Assess (Rockv)*. 2014;18(23). doi:10.3310/hta18230.

20. Sasieni P, Castanon A. NHSCSP audit of invasive cervical cancer. 2012;(July).

21. Mandelblatt JS, Lawrence WF, Womack SM, et al. Benefits and costs of using HPV testing to screen for cervical cancer (Structured abstract). *Jama*. 2002;287(18):2372-2381.

22. Elbasha EH, Dasbach EJ, Insinga RP. Model for Assessing Human Papillomavirus Vaccination Strategies. 2007;13(1):28-41.

23. Insinga RP, Glass AG, Myers ER, Rush BB. Cervical Cancer Screening : Event Duration and Health Utility Loss. 2015. doi:10.1177/0272989X07302128.

24. De Kok IMCM, van Rosmalen J, Dillner J, et al. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. *BMJ*. 2012;344:e670.

25. Simonella L, Howard K, Canfell K. A survey of population-based utility scores for cervical cancer prevention. 2014;7(1):1-11. doi:10.1186/1756-0500-7-899.

26. Goldie SJ, Kohli M, Grima D, et al. Projected clinical benefits and cost-effectiveness of a human papillomavirus 16/18 vaccine. *J Natl Cancer Inst*. 2004;96(8):604-615.

27. Krahn M, McLachlin M, Pham B, et al. Liquid-based techniques for cervical cancer screening: systematic review and cost-effectiveness analysis (Structured abstract). *Heal Technol Assess Database*. 2008;(2):79.

28. Korfage IJ, Ballegooijen M Van, Wauben B, Looman CWN, Habbema JDF, Essink-bot M. Having a Pap smear , quality of life before and after cervical screening : a questionnaire study. 2012:936-944. doi:10.1111/j.1471-0528.2012.03344.x.

CONFIDENTIAL

Appendix - Clinical impact and cost-effectiveness of primary human papilloma virus testing

A1 Parameterisation of Sexual Behaviour

Sexual debut

Entry into the HPV-susceptible population is determined by age of sexual debut. Responses from the National Survey of Sexual attitudes and Lifestyles 2010 (NATSAL-3) are used to construct the cumulative probability distributions of age at sexual debut.

The data allow us to directly determine the cumulative fraction women that are sexually active from age 16 to 75 years. For individuals aged 16 years and under, we use the distribution of reported age at first sex, S_d, (for those that report sexual activity before the age of 16 years), and scale this to the known fraction of individuals that are active by age 16 years, to determine the probability of sexual debut from age 10 onwards ($P[S_d=d] = P[S_d=d \,|\, S_d<=16]P[S_d <=16]$).

A hill function is used to provide a smooth monotone description of the empirical cumulative distribution of sexual debut age before and after the age of 16 years (shown in Figure A1.

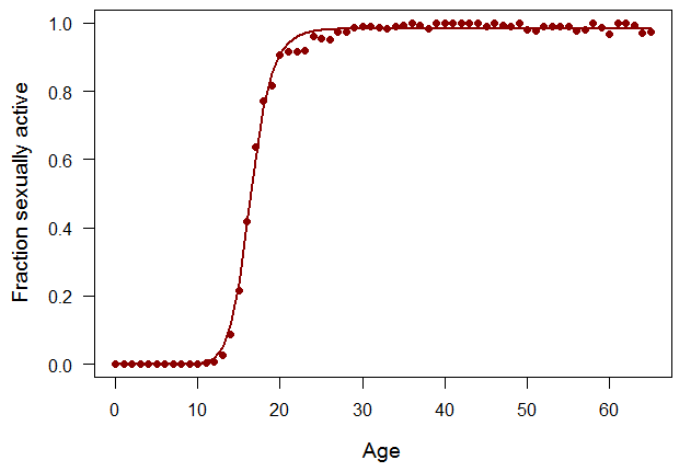


Figure A1: Cumulative probability of sexual debut as a function of age; age reported in NATSAL-3 (points) and smooth monotone function that best describes the data (line).

Partnership acquisition

The number of new partners acquired in the last year, as reported in NATSAL-3, reveals a trend towards decreasing partner acquisition with age with significant variation between individuals at the population level.

We divide the population into four sexual behavioural categories according to behaviour in the recent past; S_i, for i={1,2,3,4}, where increasing i reflects increasing sexual activity, and represent women in the 0-40th, 41-80th, 81-97th and 98-100th percentile of a given age-band, with respective to number of partners in the past five years.

We use a poisson distribution to describe the number of partners, S_N, acquired over the last 12 months.

$$P[S_N=n] = \exp(-\mu)\,\mu^n/n!$$

We divide the population into five year age bands (16-20, 21-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56 an over) and use a nelder-mead optimisation to identify the rate of partner acquisition, μ , that best describes each sub-population within ageband *a* (S_{1_a}, S_{2_a}, S_{3_a}, S_{4_a}).

An individual within the model remains associated with a given behavioural category throughout their lifetime, however, the rate of partner acquisition associated with each behavioural category decreases with age.

Partnership duration

A survival model is used, in combination with the NATSAL-3 dataset, to parameterise the cumulative probability of relationship 'survival' as a function of time and age of women at the start of relationship. In the natsal survey, individuals reported on the three most recent partners with sexual activity within the past 5 years. To counter the bias towards observing longer-relationships in a fixed-window sampling scenario, the survival of relationships is calculated using a modified Kaplan-Meier estimator that explicitly accounts for truncation (as described by Burington and colleagues[1])

$$S(x) = \prod_{y_i<x} 1 - d(y_i)/R(y_i)$$

where $d(y_i)$ represents the number of uncensored events (relationships) of length exactly y_i ;and the denominator R counts the total number of relationship events lasting more than or equal to y_i months, but excluding events that began more than y_i months before the start of the sampling window.

$$R(y_i) = \sum_j I(t_j < y_i) - \sum_j I(y_j < y_i)$$

where t_j is the time between the start of the relationship and the start of the sampling window, measured as 5 years before the date of interview according to the natsal questionnaire design, (t_j is 0 for partnerships that began after the start of the sampling period), and $I(x)$ is the indicator function, value 1 if x is true, 0 otherwise.

A partnership is defined as complete when there has been no sexual activity for 3 months. Data from the first, second and third most recent heterosexual partner is combined.

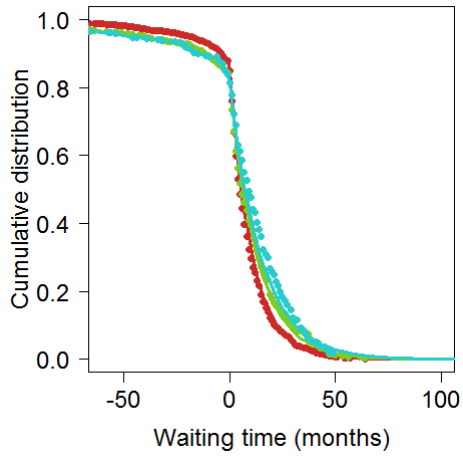
Missing partnerships: The sampling of detailed partnership information from the three most recent partnerships in the past 5 years can lead to a bias towards longer relationships and those with a large gap between relationships as they are most likely to be 'most recent' at time of interview. We compare the total number of partners in the past 5 years to the number of partners for whom we have detailed information for individual, to determine the number of missing partnerships in our data sample. It is assumed that missing partnerships are complete and therefore will be most similar to completed partnerships that have been reported in detail by the same individual.

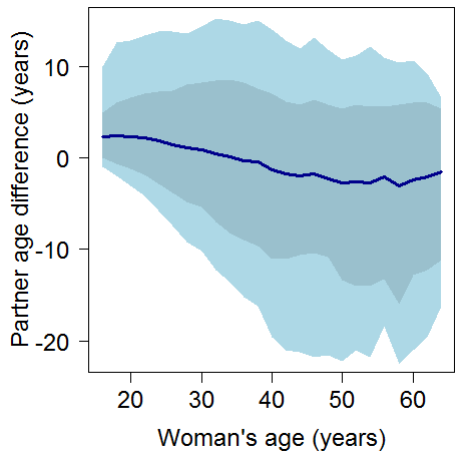
The detailed information on complete partnerships for individual *i* can be weighted by W, $W = 1 + (H_{5,i} - T_i)/C_i)$, where $H_{5,i}$ is the number of partnerships reported in the last five years; T_i is the number of partners for whom detailed partnership data is available; C_i is the number of complete partnerships for whom data is available for subject i.

Relationship survival curves exhibit a biphasic decay; with a large number of short term relationships and a smaller number of very long partnerships. The data reveals that the fraction of relationships falling into the short-term category increases as a function of age at start of relationship; the five year survival for a relationship is 19%, 7% and 3,5% for women aged 16 to 20 years, 31 to 35 years, and 51 to 55 years, respectively, at the start of relationship.

Age mixing

An age-mixing matrix is generated by directly sampling from the age of most recent partners reported by female participants in NATSAL-3. We stratify the data according to the age of female respondent at the start of relationship to reveal an increasing variance in partner age for older women. This approach is preferred to a more traditional approach to partner-matching that assumes a constant age difference distribution or an approximation of +/- 3 years, as it better captures the complexities of HPV transmission; in particular, the role of novel HPV infections in older women versus long term persistent infection. The shift in age difference is illustrated in Figure A2.





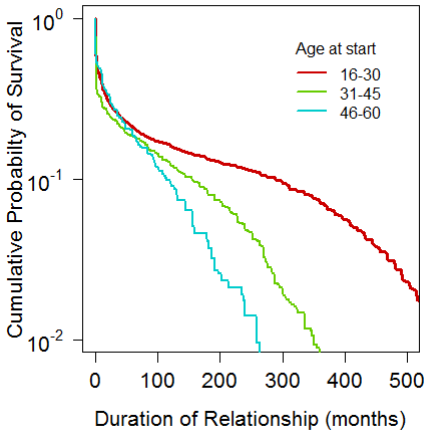


Figure A2: Sexual behaviour parameters are identified using data collected in natsal-3. (A) Simulated cumulative probability of a waiting time to next partner after the end of previous relationship; (B) cumulative probability relationship survival as a function of time since start of relatiosnhip; (C) age preference of male partners as a function of female age at start of relationship.

Frequency of sex acts

The model simulates HPV transmission by introducing a probability of transmission per sexual contact. We use data from NATSAL-3 to quantify the number of sex acts per month for individuals in an active relationship (defined as those participants who reported a sexual encounter in the last 3 months with the most recent partner). As above, a weighting is added to response data that scales with total 5 year partner count to remove the bias towards reporting of characteristics from long term relationships (Figure A3).

We also derive the fraction of new relationships that result in a single sexual contact only. We distinguish between recent ongoing relationships that may yet lead to further contacts and those that are complete by assuming that relationship is complete if there have been no contact in the past three months.

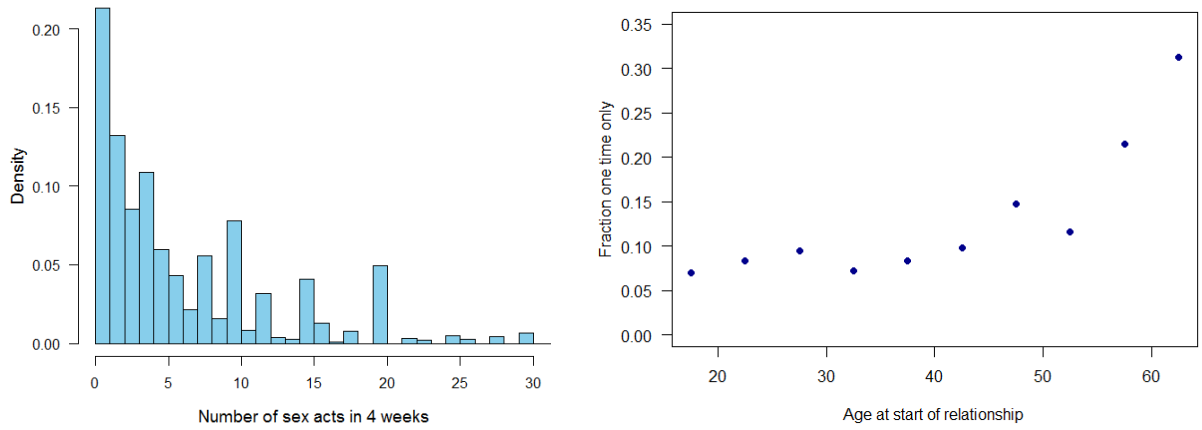


Figure A3: Left panel: Distribution of reported sex acts in a 4 week period for individuals in an active relationship. Right panel: Fraction of new partnerships that result in a single contact only.

A2 Parameterisation of HPV infection

The sexual behaviour characteristics described above are used to generate the age of sexual debut and formation and dissolution of subsequent sexual partnerships for a population of 100,000 women from birth to age sixty five years.

The HPV status of a new male partner is randomly generated using the distribution of age of new partners, reported in NATSAL-3, according to the age of the woman at start of partnership, and the prevalence of HPV among men of the preferred partner age.

HPV infection is modelled by introducing a per-sex-act probability of transmission of HPV. The probability of infection by each strain is assumed to be independent. In this work we consider HPV-16, 18, 31, 33, 45, 51 and 52. Rates of clearance and transmission of HPV were parameterised using HPV prevalence data in women and sero-prevalence data in males.

HPV-strain specific prevalence was determined using surveillance data collected by Public Health England from residual samples taken from the NHSCSP pre-immunisation

for women aged 25-65 years[2]. For younger women between the ages of 16 and 24 years, HPV prevalence was measured in residual samples taken from the national chlamydia screening programme (NCSP), pre-HPV-immunisation[3]. Data from these younger women is important for characterising the peak of HPV infection, however, the selective nature of women attending the NCSP means that the sample reflects a higher sexual risk group than the general population; NCSP data is accompanied by data with number of partners reported in the past 12 months which is higher than that predicted by NATSAL-3 for women aged16-24 years. We introduce a weighting to resample the NCSP data such that the number of partners reported in the past 12 months matches that observed in natsal-3 for women aged 16-24 years, and introduce a sub-population of sexually-inactive women in the same age range (as predicted in NATSAL-3 but not present in the NCSP dataset) who are expected to be HPV-naïve. HPV prevalence is recalculated in this re-weighted population and this new comparable prevalence is merged with the NHSCSP predicted prevalence.

HPV prevalence data was not available for a sufficiently large population in England. Instead, national surveillance data describing male sero-prevalence of HPV-16 and HPV-18 in England, collected by PHE[4], was used to estimate prevalence of HPV among males. A study of sero-prevalence in the Netherlands revealed that sero-prevalence levels were similar in HPV-,33, 45 and 52, but approximately two-fold higher in HPV-31[5] ; a similar result was found in the German population[6]. In the parameterisation that follows, sero-prevalence of hpv 33, 45 and 52 among males was constrained by observed sero-prevalence of hpv-18, in accord with levels of hpv prevalence of these strains observed in women in England. Sero-prevelance of HPV strains 31 and 51 was estimates by scaling the observed hpv-18 sero-prevalence according to the ratio of hpv prevalence of hpv-31: hpv-18 and hpv-51: hpv-18 observed in women.

HPV infection in males

A static model of transmission is applied in which male prevalence is assumed to be constant throughout the duration of these simulation; we argue that the introduction of primary HPV DNA testing in cervical screening will not have an effect on male prevalence of HPV.

A simple three compartment differential equation model is used to analyse the sero-prevalence data and extract HPV prevalence for each HPV type. We consider individuals that are (i) infected but sero-negative (I); (ii) sero-positive, that is they have detectable HPV antibodies (S), and (iii) infected and HPV-DNA positive (H).

$$\dot{I} = f(t) - (c + k)I(t)$$

$$\dot{S} = k\, I(t) - wS(t)$$

$$\dot{H} = f(t) - cH(t)$$

Where, *f(t)* is the number of new infections at time *t*; *c* is the rate of clearance of male infection; *k* is the rate of sero-conversion; and *w* is the rate of HPV antibody waning.

The size of the infected population, H, can be estimated using numerical methods to solve the following equation:

$$\dot{H} = \frac{\ddot{S}}{k} + (w + c + k)\frac{\dot{S}}{k} + w(c + k)\frac{S}{k} - cH(t)$$

where, the observed sero-prevalence, S[t], can be described by a polynomial, and it is assumed that the half life of antibodies is at least 20 years, that is the rate of waning (w) is constrained to be less than 0.05 (/year).

The rates of sero-conversion and clearance for each male HPV strain are identified, together with the rates of female clearance and transmission, using the observed HPV sero-prevalence in males and prevalence in females. Described in detail below.

Parameterisation

The disease transmission was parameterised independently for each HPV-subtype. In this parameterisation we assume that HPV prevalence is not sensitive to screening strategy. The reasoning is that (i) the number of women treated for cervical lesions is small relative to the number of women that are infected with HR-HPV, ~10% of population at large; and (ii) not all treatment of lesions leads to clearance of HPV-infection. As a result, we can identify the rate of disease transmission and clearance using a simplified individual-based model without screening, in a computationally tractable parameterisation.

A Markov Chain Monte Carlo simulation, using an adaptive Metropolis algorithm, was implemented using the FME package in R to simultaneously identify the (i) HPV clearance parameters in females and males (c_1, c_2, c_m); (ii) per contact probability of transmission from males to females; and (iii) rate of sero-coversion in males. Each chain was run for a length of 20,000 and 100 parallel chains were generated using randomly generated starting values, for each strain of HPV.

A thinning interval of 50 was used to remove auto-correlation within each chain. Convergence was identified using the Geweke test statistic, a test of equality of the means of the first 10% and last 50% of the markov chain. A Gelman convergence diagnostic was then used to confirm convergence of the MCMC output in the parallel chains; a comparison of the empirical variance of each parameter within each chain should be comparable to the variance from all chains combined. The final parameter distribution reflects the joint distribution of the parallel chains.

Clearance of HPV infection

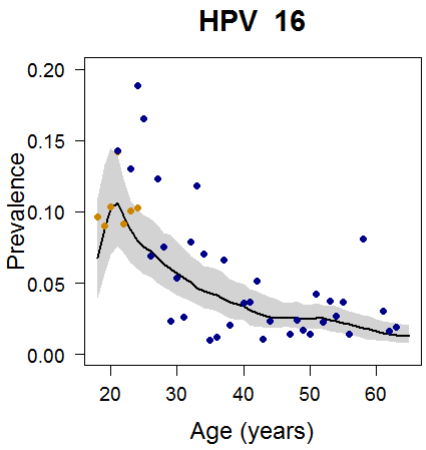
We model the waiting time to clearance of infection using a weibull distribution C~W[c_1, c_2] , where a c_1 value of less than 1 gives a decreasing rate of clearance with time and determines the scale of the distribution. We find that all HPV types are well described by a decreasing rate of clearance with increasing time since infection, that is the shape parameter lies below 1. The scale parameters lead to the differences in the time of clearance, and the analysis suggests that HPV types 18 and 51 are cleared most rapidly, with just 16% and 25% of women expected to remain infected for 12 months, respectively, compared to 54% and 69% of women infected with HPV types 16 and 31

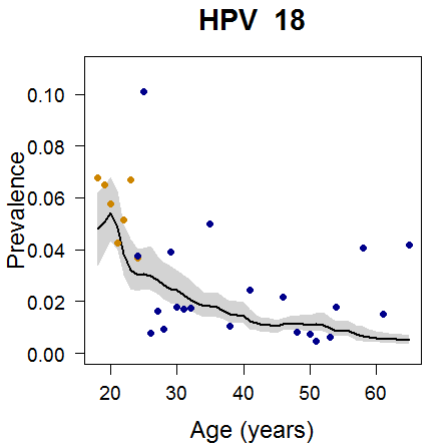
respectively. The most persistent long-term infections are found to be associated with hpv –strains 31, 33 and 45.

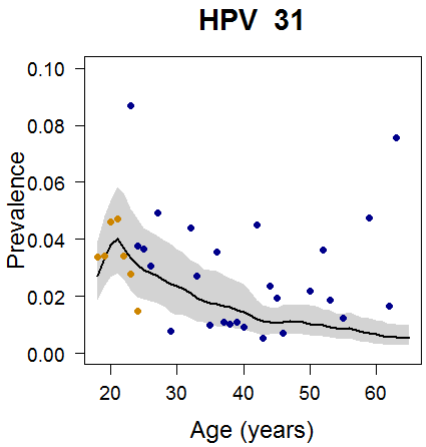
The rate of male clearance is not well constrained, but the 95% confidence interval of rates suggest a half-life of male infection of at most 7 months across all HPV types. Under the assumption that HPV antibodies waning results in a half-life of at least 20 years, the model predicts low rates of male sero-conversion of 0.043 (/year) for HPV-16 and 0.011 (/year) for HPV-18.

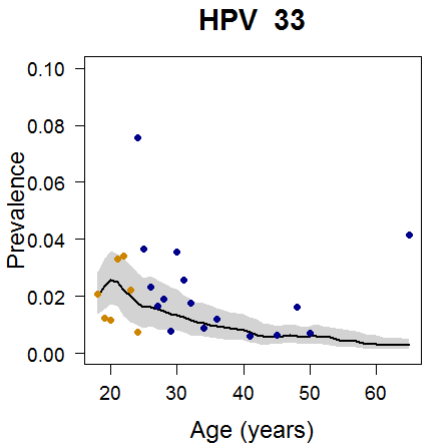
HPV transmission probability (per-sex act)

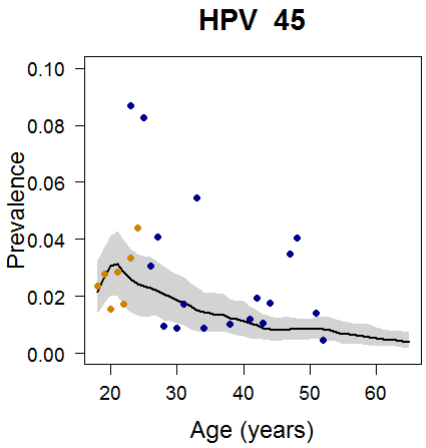
We find that the transmission probability per contact is not well defined. One explanation for this is that, according to the sexual behaviour data, the majority of partnerships result in multiple contacts; where the probability of contracting HPV from an infected partner, 1-(1- *Transmission.Probability*)^*'Sex.Acts'*, becomes decreasingly sensitive to the *Transmission.Probability* as the number of Sex.Acts increases. We accept the broad range of values suggested for transmission probabilities as they are able to reproduce the observed HPV prevalence within the context of known sexual behaviour.

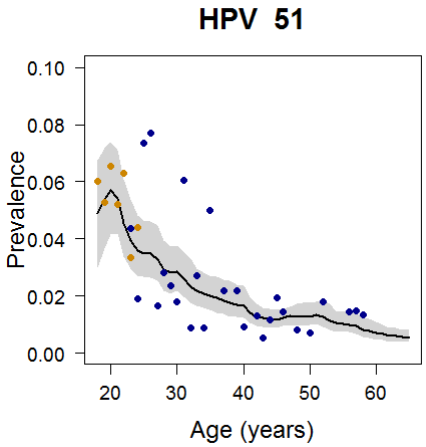












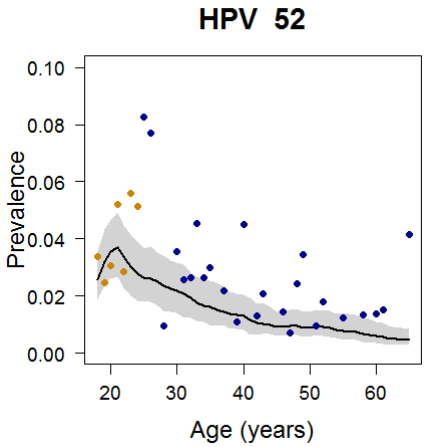
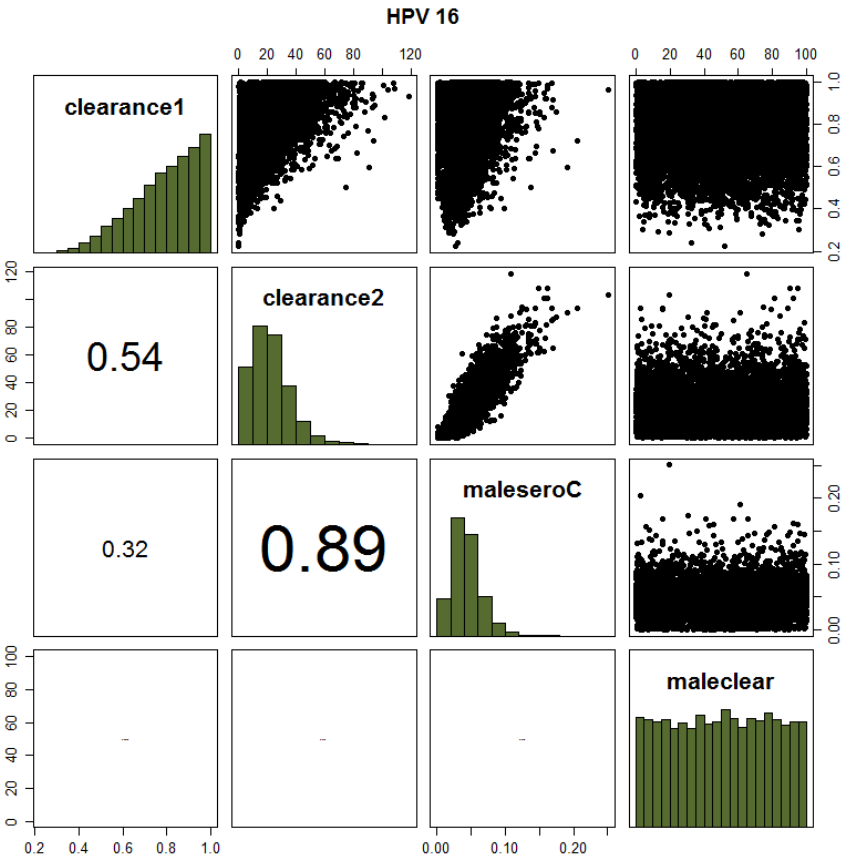


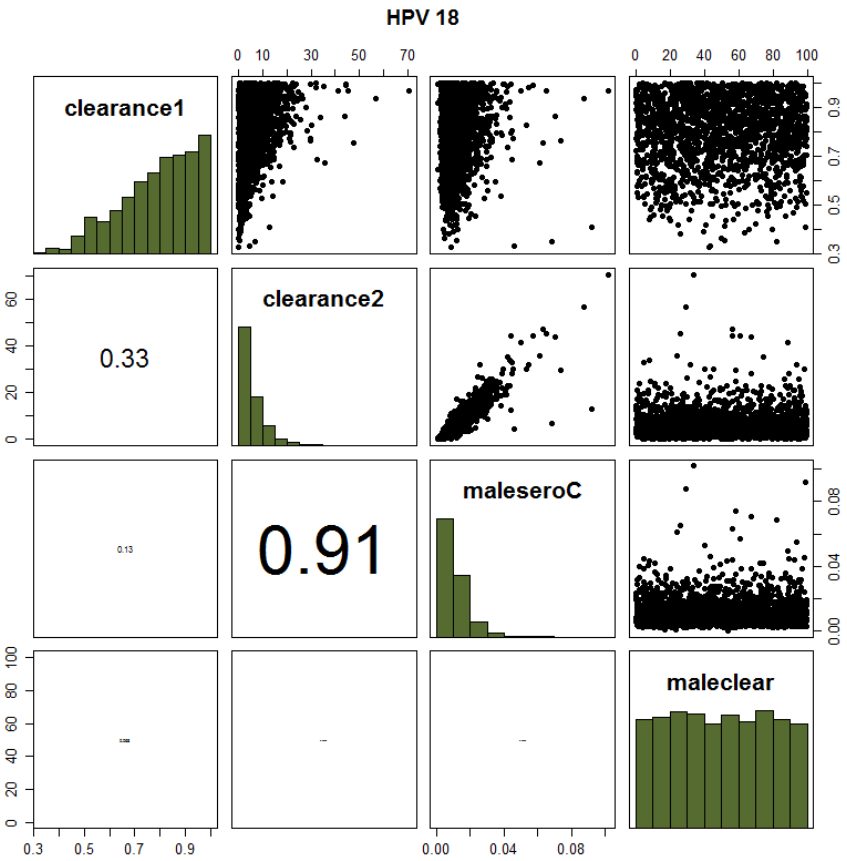
Figure A4: Observed type-specific HPV prevalence measured in residual samples from the NCSP (age 16-24 years – orange points) and NHSCSP (age 24-65 years – blue points) and best-fitting model predictions – mean (solid black line), upper and lower 95% interval (grey shaded region).

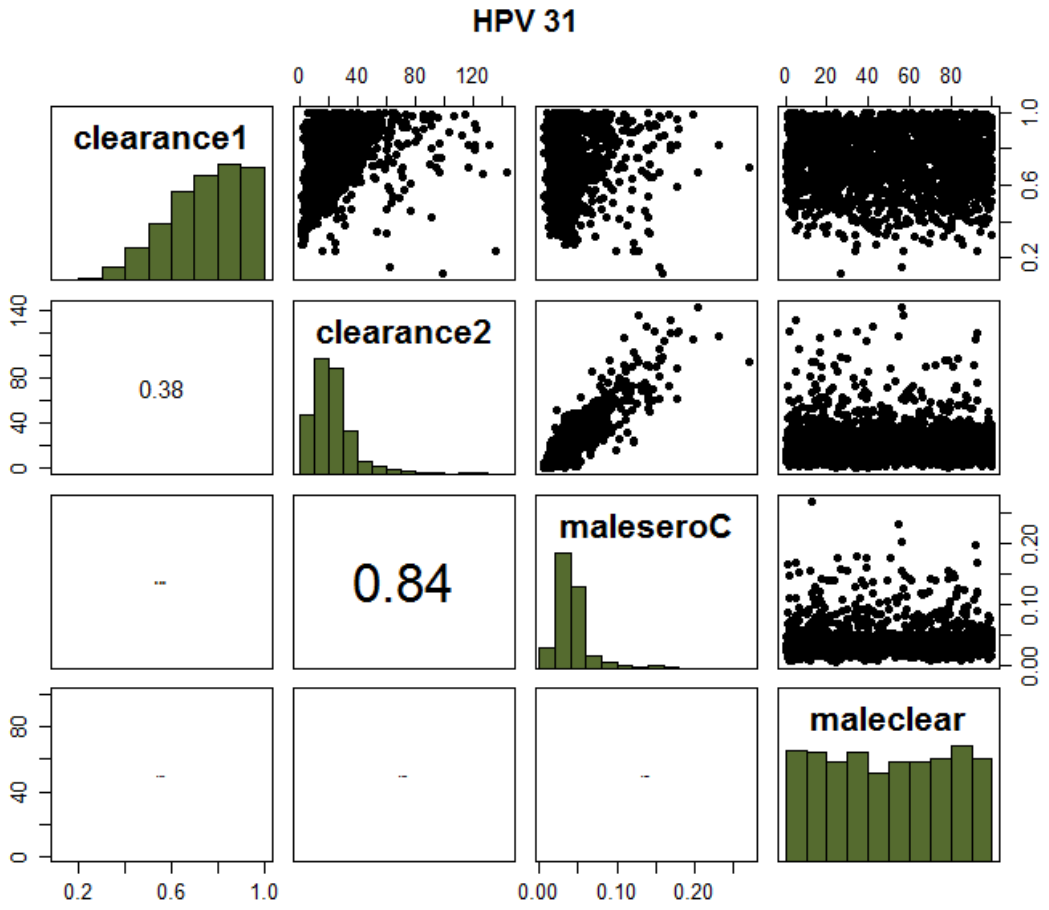
	Per-contact probability of transmission	c1	c2	Fraction of women remain infected at 12 months	Fraction of women remain infected at 24 months	Male seroconversion (/year)	Male rate of clearance (/year)
HPV 16	0.536	0.802	22.274	0.495	0.321	0.042	51.001
	(0.045 ,0.98)	(0.459 ,0.993)	(1.183 ,53.791)	(0.007 ,0.778)	(0 ,0.614)	(0.002 ,0.09)	(1.975 ,97.415)
HPV 18	0.580	0.790	5.259	0.137	0.051	0.010	49.922
	(0.051 ,0.985)	(0.475 ,0.991)	(0.223 ,17.243)	(0 ,0.484)	(0 ,0.271)	(0.003 ,0.03)	(2.61 ,96.616)
HPV 31	0.452	0.735	18.404	0.458	0.287	0.035	50.609
	(0.023 ,0.95)	(0.395 ,0.983)	(2.756 ,37.403)	(0.15 ,0.698)	(0.058 ,0.503)	(0.014 ,0.05)	(2.538 ,97.065)
HPV 33	0.422	0.716	10.116	0.304	0.155	0.037	51.348
	(0.012 ,0.966)	(0.363 ,0.986)	(1.18 ,22.975)	(0.046 ,0.568)	(0.006 ,0.355)	(0.014 ,0.049)	(2.776 ,96.175)
HPV 39	0.499	0.786	14.462	0.374	0.209	0.021	50.820
	(0.027 ,0.978)	(0.47 ,0.992)	(1.354 ,38.19)	(0.023 ,0.702)	(0.002 ,0.513)	(0.005 ,0.045)	(3.367 ,97.786)
HPV 45	0.418	0.728	14.340	0.388	0.226	0.036	49.239
	(0.019 ,0.974)	(0.404 ,0.987)	(1.625 ,31.474)	(0.049 ,0.645)	(0.008 ,0.452)	(0.01 ,0.05)	(3.561 ,97.703)
HPV 51	0.518	0.789	7.630	0.205	0.093	0.012	50.723
	(0.042 ,0.975)	(0.47 ,0.99)	(0.272 ,26.266)	(0 ,0.606)	(0 ,0.397)	(0.002 ,0.032)	(2.848 ,98.391)
HPV 52	0.454	0.764	17.917	0.460	0.279	0.037	48.677
	(0.024 ,0.97)	(0.428 ,0.99)	(3.463 ,34.4)	(0.163 ,0.683)	(0.057 ,0.485)	(0.016 ,0.049)	(2.464 ,97.512)
HPV 58	0.486	0.776	13.805	0.355	0.200	0.020	49.638
	(0.037 ,0.965)	(0.43 ,0.99)	(1.047 ,37.687)	(0.007 ,0.707)	(0 ,0.509)	(0.004 ,0.044)	(3.281 ,96.428)
HPV 59	0.529	0.814	6.993	0.195	0.074	0.011	50.930
	(0.041 ,0.981)	(0.48 ,0.993)	(0.43 ,19.868)	(0 ,0.526)	(0 ,0.305)	(0.003 ,0.025)	(2.452 ,97.32)
HPV 66	0.510	0.789	8.938	0.243	0.113	0.012	48.510
	(0.04 ,0.985)	(0.445 ,0.994)	(0.471 ,29.398)	(0 ,0.624)	(0 ,0.422)	(0.003 ,0.036)	(1.951 ,97.546)

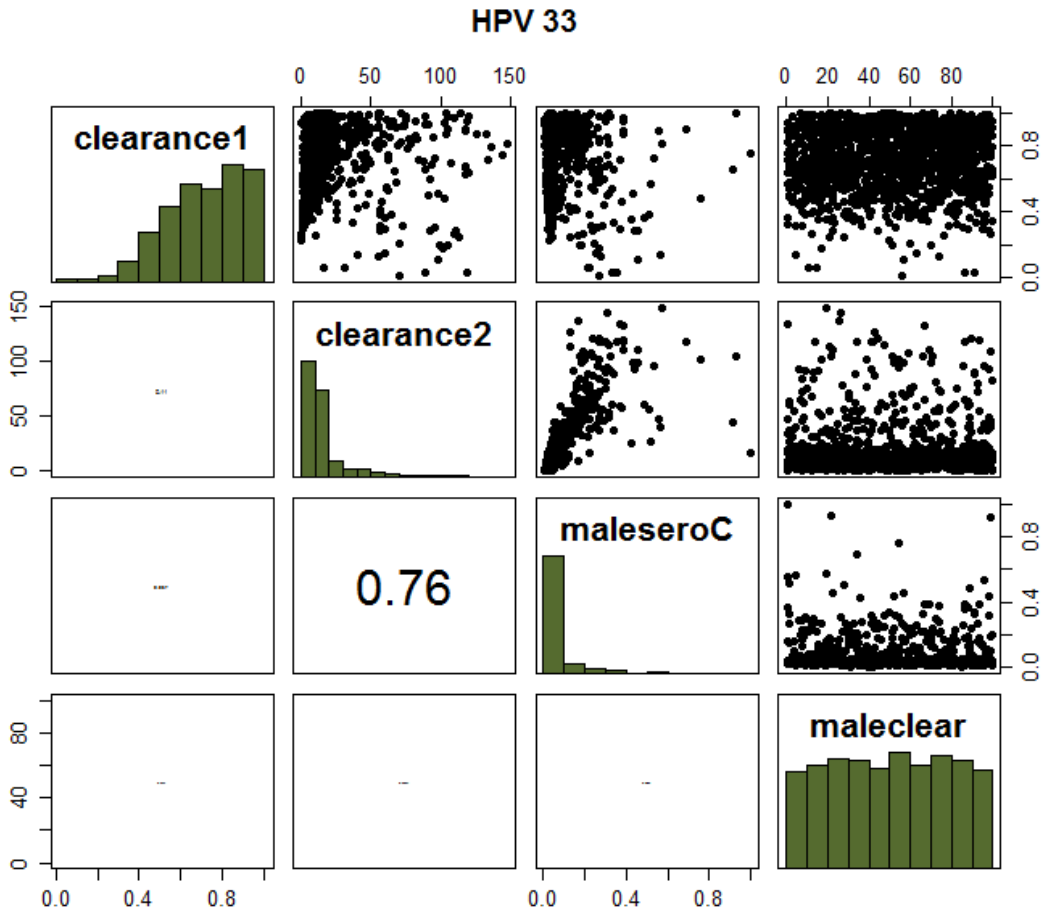
Table A1: Best fitting clearance and transmission parameters for HPV-16, 18, 31, 33, 39, 45, 51, 52, 58, 59 and 66

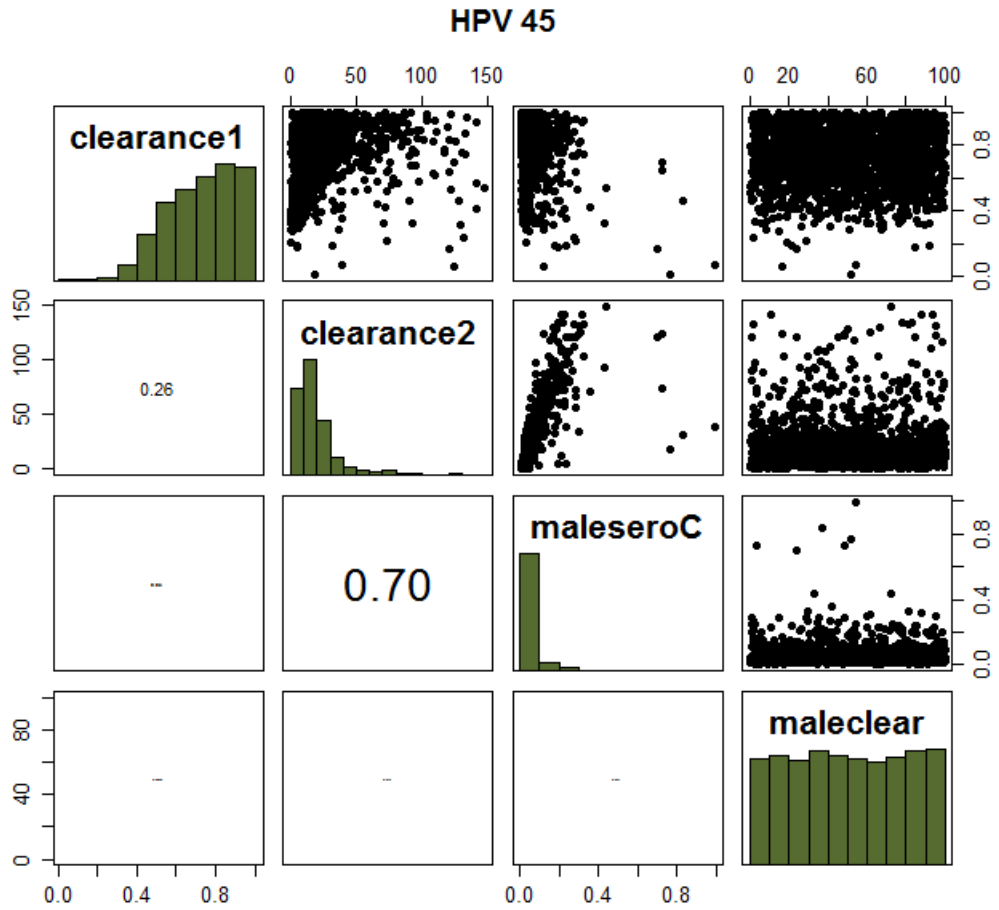
Figure A5: Posterior distributions of best fitting clearance and transmission parameters for HPV-16, 18, 31, 33, 45, 51 and 52

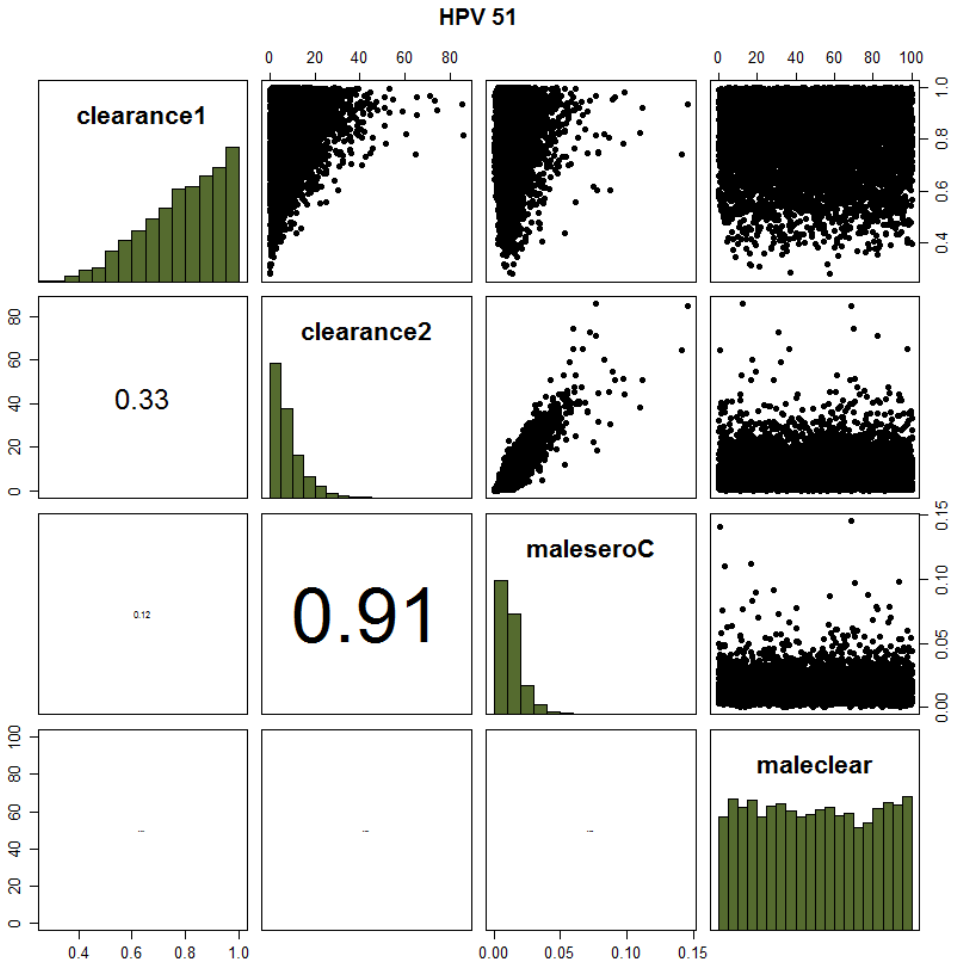


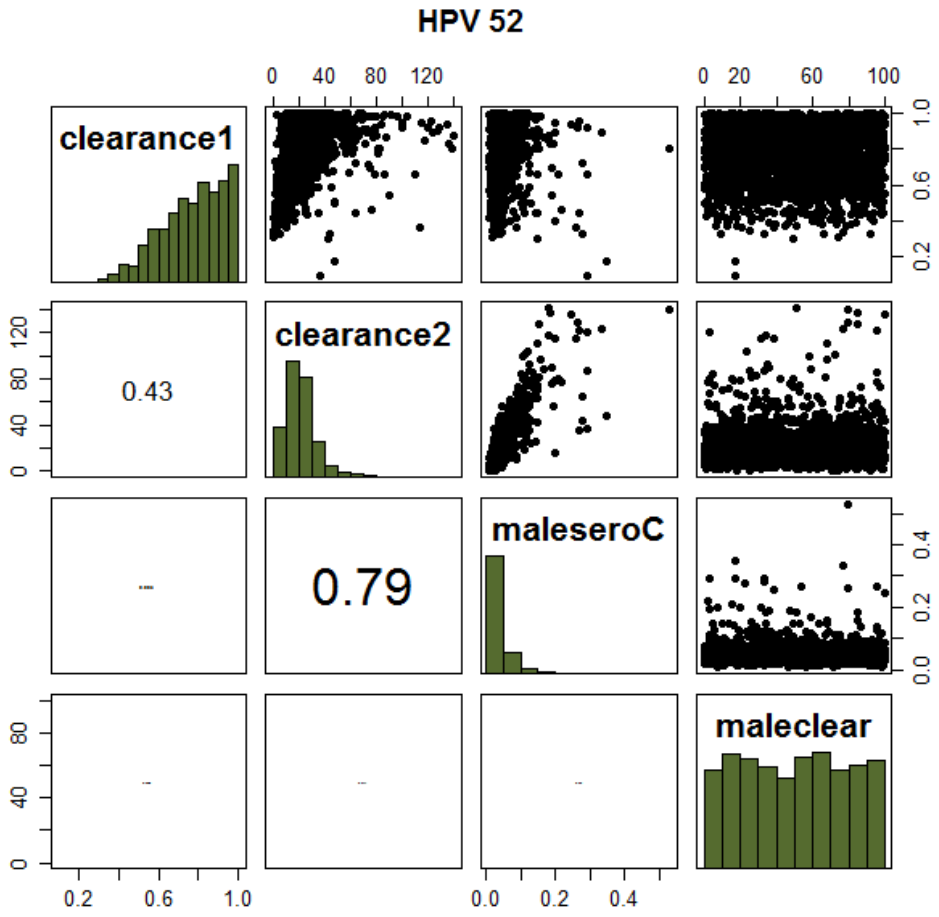



HPV 31
clearance1
clearance2
maleseroC
maleclear
0.38
0.84


HPV 33
clearance1
clearance2
maleseroC
maleclear
0.76


HPV 45
clearance1
clearance2
maleseroC
maleclear
0.26
0.70


HPV 51
clearance1
clearance2
maleseroC
maleclear
0.33
0.12
0.91


HPV 52
clearance1
clearance2
maleseroC
maleclear
0.43
0.79
0 40 80 120
0 20 60 100
0 40 80 120
0.2 0.6 1.0
0.0 0.2 0.4
0 40 80
0.2 0.6 1.0
0.0 0.2 0.4

A3 Parameterisation of Disease Progression

The disease progression was parameterised using surveillance data collected by PHE; residual LBC samples from the NHSCSP were used to identify cytological outcome, is measured as a function of age and hpv strain. We assume that each infection within an individual will progress according to the hpv strain and time since infection. Multiple HPV infections can lead the development of multiple lesions with independent progression rates. We explicitly model cytological outcomes, rather than discrete CIN states, in order to directly calibrate the model against observed cytology outcomes. The advantage of this approach is that we avoid incorporating assumptions about the sensitivity and specificity of colposcopy and cytological testing into the underlying model parameters.

We use a nested conditional probability structure to generate a model in which the probability of distinct cytological outcomes varies as a continuous function of time since infection for each strain of HPV, rather than distinct disease states.

We introduce a flexible mixed exponential model structure, in which the probability of a normal cytological outcome can decrease and then increase as a function of time post-infection (T), according to choice of parameters p_norm1 and p_norm2:

$$P[N|T=t] = 1- (1-Exp(-p\_norm1*t))Exp(-p\_norm2*t)$$

For subsequent cytological outcomes, we define a structure that results in an increasing probability of a severe outcome with time. Given an abnormality, we model the probability of a borderline outcome as a decreasing function of time since infection,

$$P[B|\ Norm^c, T=t] = Exp(-p\_bord*t)$$

Similarly, the probability of a mild and moderate outcomes given that the outcome was not normal nor borderline, and, not normal nor borderline nor mild, respectively, is described by:

$$P[Mild|\ Norm^c \text{ and } Bord^c, T=t] = Exp(-p\_mild*t)$$

$$P[Mod|\ Norm^c \text{ and } Bord^c \text{ and } Mild^c, T=t] = Exp(-p\_mod*t)$$

Finally, the probability of a severe outcome is modelled as:

$$P[Sev] = 1 – P[Norm] – P[Bord]- P[Mild]-P[Mod]$$

This leads to waves of disease progression with time. The model was fitted simultaneously for each HPV-strain, in a simulation that incorporated screening under the existing primary cytology protocol.

A Markov Chain Monte Carlo simulation, using an adaptive Metropolis algorithm, was implemented using the FME package in R to simultaneously identify the rate of change in probability of a given cytology outcome with time. Each chain was run for a length of 20,000. Clearance and transmission parameters were sampled from the posterior distributions derived previously for each HPV type; 200 distinct combinations were used in total. 50 parallel chains were generated for each clearance-transmission parameter-combination using randomly generated starting values.

As before, a thinning interval of 50 was used to remove auto-correlation within each chain. Convergence was identified using the Geweke test statistic, a test of equality of the means of the first 10% and last 50% of the markov chain. A Gelman convergence diagnostic was then used to confirm convergence of the MCMC output in the parallel chains; a comparison of the empirical variance of each parameter within each chain should be comparable to the variance from all chains combined. The final parameter distributions reflect the joint distribution of the parallel chains generated using all 200 clearance parameter-combinations.

Cancer progression

Cervical lesions can be dissected into sections of different grades, each infected with a unique high risk HPV strain, suggesting that HPV infection with different strains can lead the development of multiple lesions with independent growth rates. HPV typing was carried out on residual tissue sections from routinely obtained diagnostic biopsies of cervical cancers archived in NHS pathology laboratories [n=555] by Howell-Jones and colleagues[2]. The observed distribution of HR-HPV types and co-infections in this sample was used to estimate the number of adenocarcinomas and squamous cell carcinomas that are associated with each strain, nationally.

Co-infection of cancer–causing strains, as defined in our model, was observed in ~7.1%, and 10%, of tissue samples taken from cervical cancers, and adenocarcinomas, respectively; ~3%, and 6%, of samples were positive for both hpv-16 and hpv-18 in cervical cancers, and adenocarcinomas, respectively. We generate national cancer incidence for each model HPV-type alone plus co infection of hpv 16 and hpv 18 by scaling the incidence values with the observed distribution of types. There are not sufficient data to accurately project the co-infection with other strains; instead, we distribute the remaining joint infection cases according to the number of observed cancers associated with a single infection of each type involved.

Adenocarcinoma and Squamous cell carcinoma

The progression of women to adenocarcinoma is modelled independently form squamous cell carcinoma. It is assumed that both conditions might arise independently. We identify the waiting time to both types of invasive cancer using reported cancer incidence in combination with data form the cervical cancer audit identifying the distribution of adenocarcinoma and squamous cell carcinoma according to age. The hazard of both squamous cell carcinoma and adenocarcinoma incidence is assumed to increase as a function of time post-infection, and we model the waiting times for each hpv type using a gamma distribution. We assume that 75.9% of cancer cases are diagnosed in FIGO stage 1, and the remainder are assumed to be stage 2+ [source: audit of invasive cervical cancer, July 2011].

Age-dependent cancer survival rates

Cancer mortality rates are calculated using 1 and 5 years survival rates published by ONS. We describe the survival using an exponential decay function following diagnosis of cancer and estimate an age-dependent mortality hazard inTable . Rates are identified using a nelder-mead optimisation in R.

CONFIDENTIAL

Age at diagnosis	Rate of mortality
15-39	0.03
40-49	0.05
50-59	0.10
60-69	0.14
70-79	0.28
80-99	0.50

Table A2: Age dependent mortality rate following diagnosis of cervical cancer

Table A3: Model parameters that best describe the occurance of cytological abnormalities with time since infection.

	p.norm1	p.norm2	Expected wait to abnormality (years)	p.bord	p.mild	p.mod
HPV 16	0.334 (0.08, 1.574)	0.001 (0, 0.004)	3.0	0.723 (0.091, 1.907)	0.658 (0.046, 1.912)	0.737 (0.048, 1.906)
HPV 18	1.381 (0.233, 1.982)	0.005 (0, 0.038)	0.7	1.098 (0.13, 1.944)	0.848 (0.045, 1.939)	0.954 (0.057, 1.939)
HPV 31	0.187 (0.018, 0.938)	0 (0, 0.001)	5.3	0.575 (0.03, 1.911)	0.606 (0.023, 1.893)	0.845 (0.037, 1.948)
HPV 33	0.421 (0.121, 1.671)	0 (0, 0.003)	2.4	0.451 (0.103, 1.685)	0.123 (0.021, 0.491)	0.838 (0.065, 1.945)
HPV 45	0.15 (0.006, 0.203)	0.001 (0, 0.007)	6.6	0.91 (0.039, 1.928)	0.86 (0.031, 1.945)	0.937 (0.047, 1.939)
HPV 51	0.998 (0.126, 1.965)	0.004 (0, 0.039)	1.0	0.827 (0.062, 1.87)	0.723 (0.026, 1.887)	0.933 (0.046, 1.953)
HPV 52	0.144 (0.028, 0.43)	0.001 (0, 0.005)	6.9	0.678 (0.039, 1.901)	0.677 (0.021, 1.916)	0.841 (0.036, 1.942)

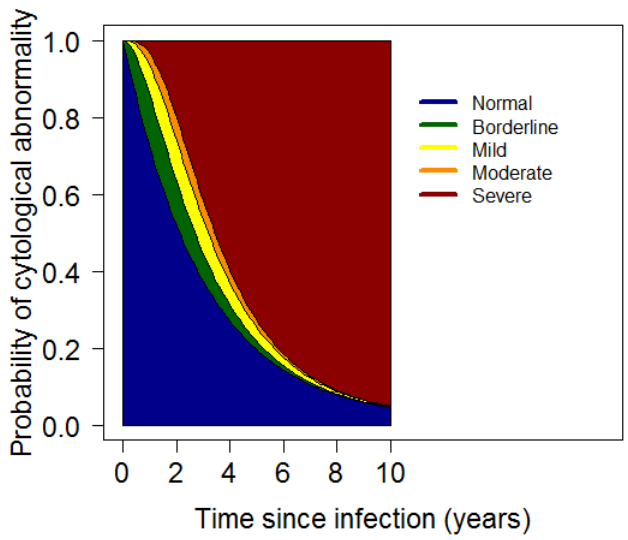
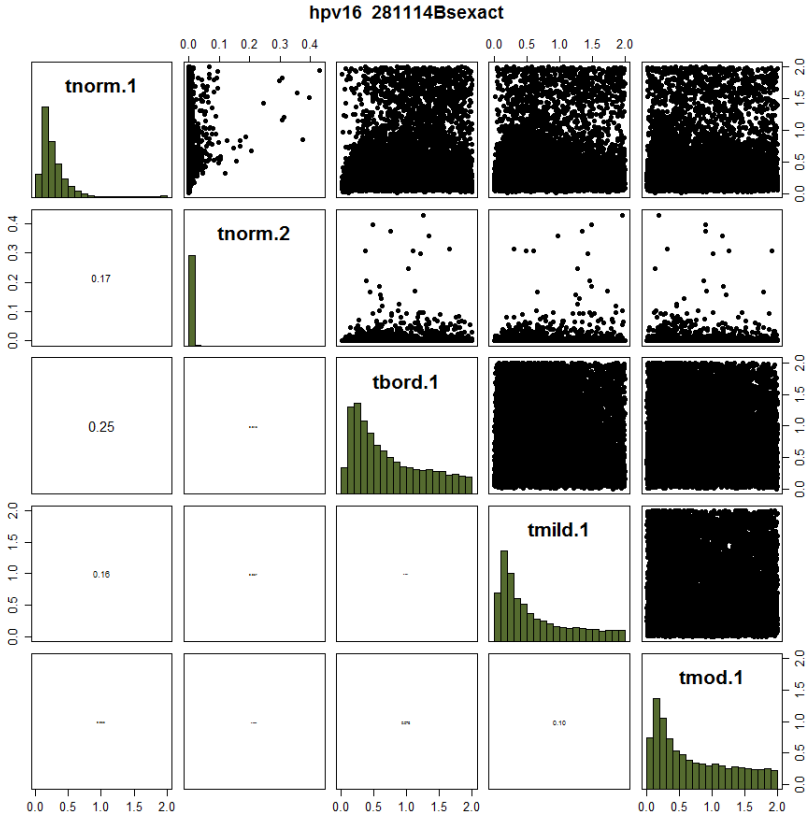
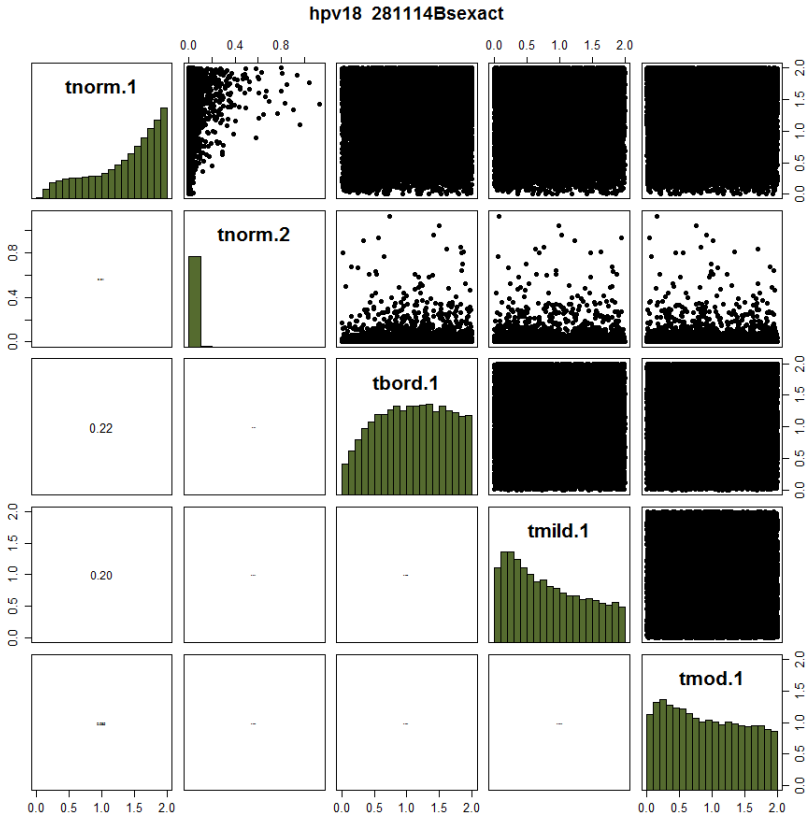
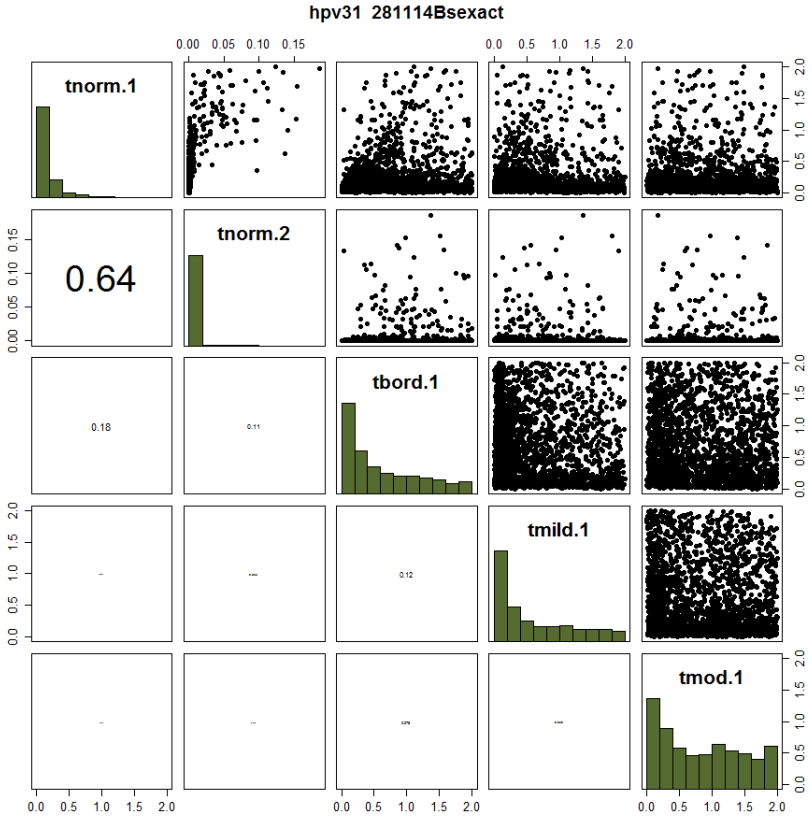


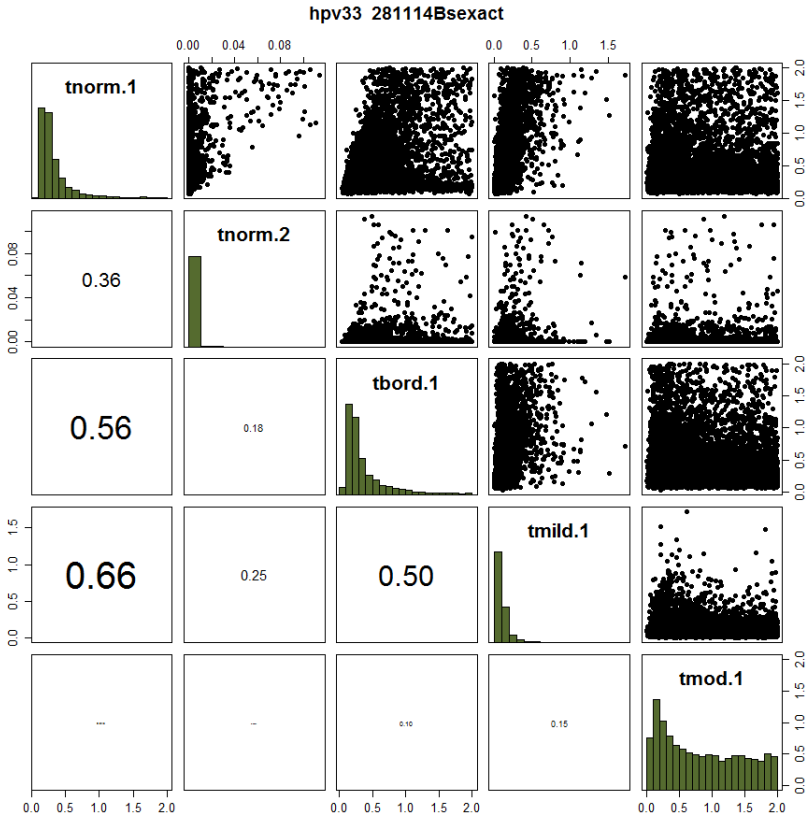
Figure A6: Illustration of probability of cytological abnormality as a function of time post- infection for HPV-16. This coincides with an decreasing probability of natural clearance with time post-infection. The time dependent probabilities are calibrated for each HPV strain using observed cytology outcome and HPV status measured as a function of age in residual samples collected by the NHSCSP.

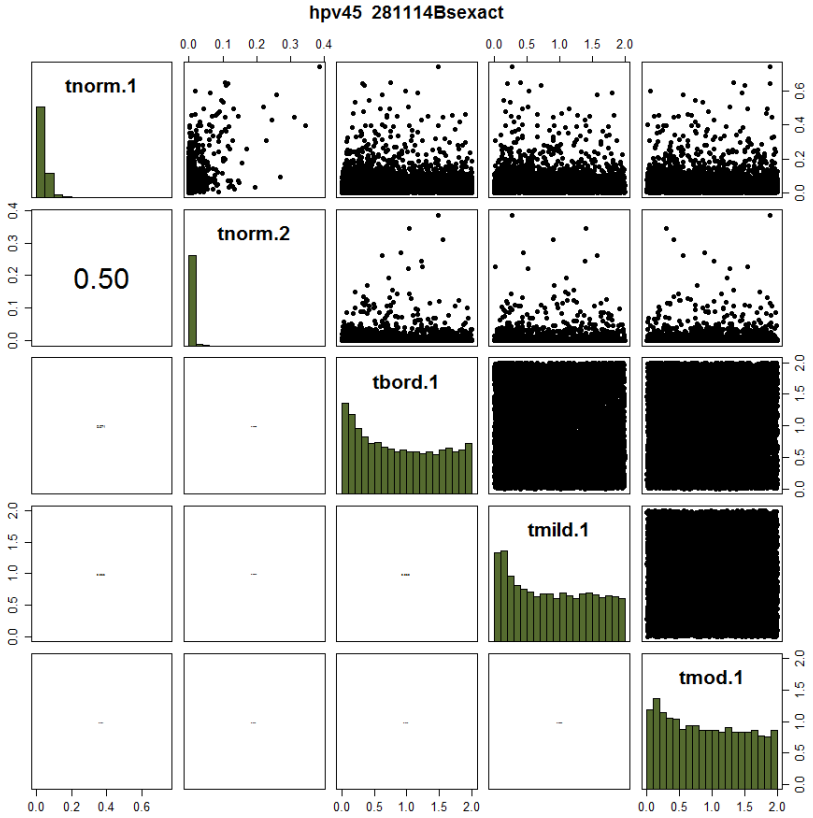
Figure A7: Posterior distributions of best fitting parameters describing cytology outcomes as a function of times for HPV-16, 18, 31, 33, 45, 51 and 52

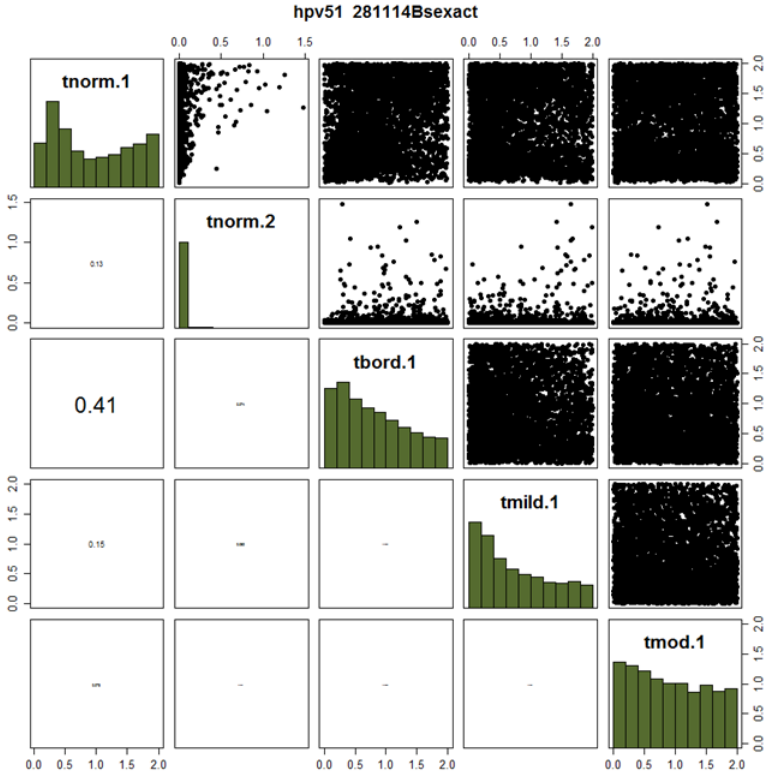

hpv16 281114Bsexact
tnorm.1
tnorm.2
tbord.1
tmild.1
tmod.1
0.17
0.25
0.16
0.10

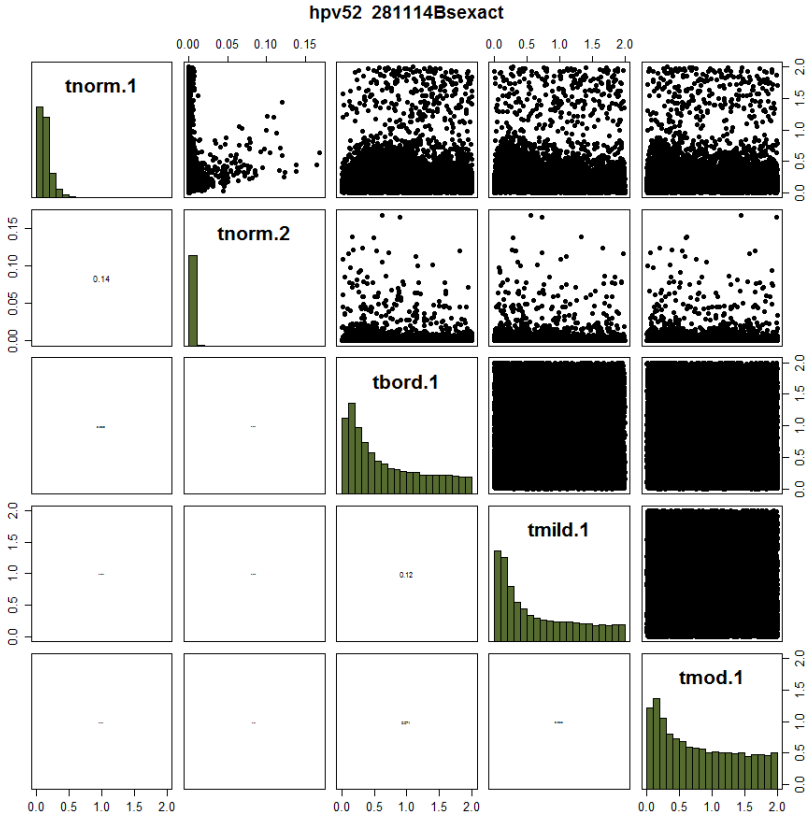

hpv18 281114Bsexact
tnorm.1
tnorm.2
tbord.1
tmild.1
tmod.1
0.22
0.20


hpv31 281114Bsexact
tnorm.1
tnorm.2
tbord.1
tmild.1
tmod.1
0.64
0.18
0.11
0.12


hpv33 281114Bsexact
tnorm.1
tnorm.2
tbord.1
tmild.1
tmod.1
0.36
0.56
0.18
0.66
0.25
0.50
0.10
0.15


hpv45 281114Bsexact
tnorm.1
tnorm.2
0.50
tbord.1
tmild.1
tmod.1


hpv51 281114Bsexact
tnorm.1
tnorm.2
0.41
tbord.1
tmild.1
tmod.1



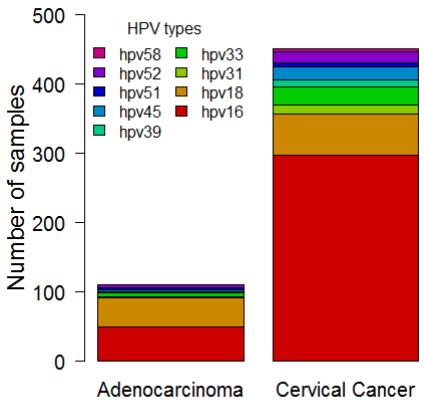


Figure A8: Hpv-type breakdown of observed adenocarcinomas and squamous cell carcinomas.

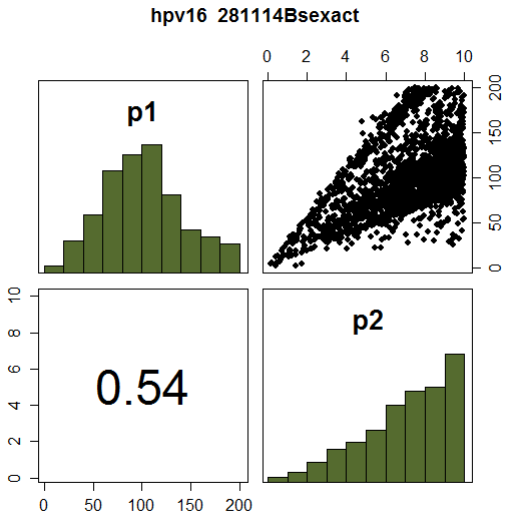
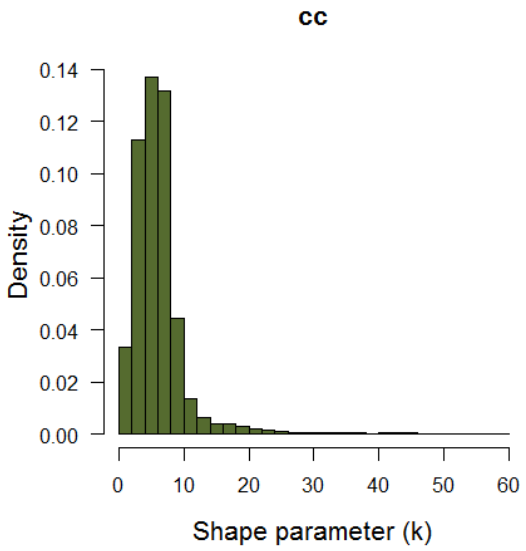
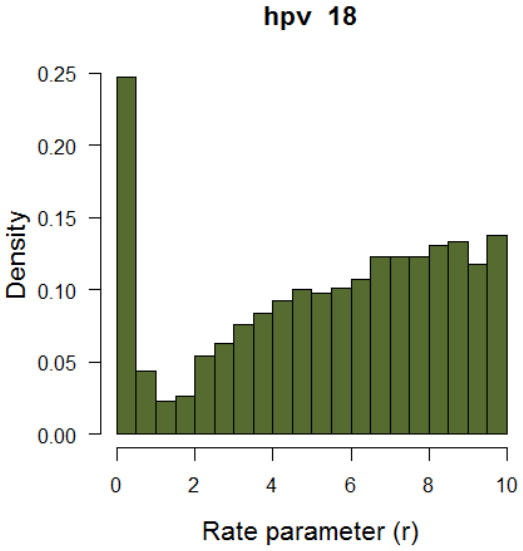
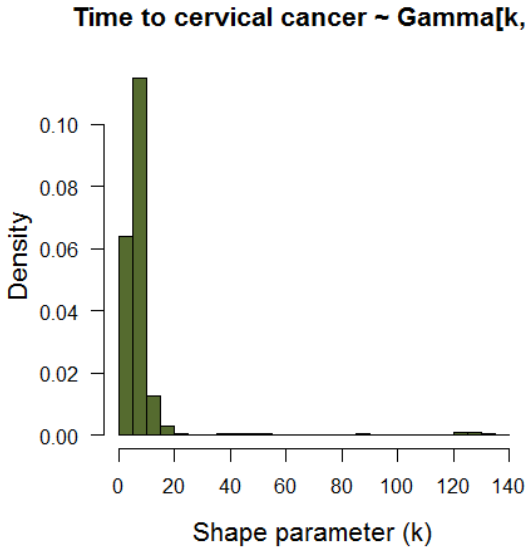
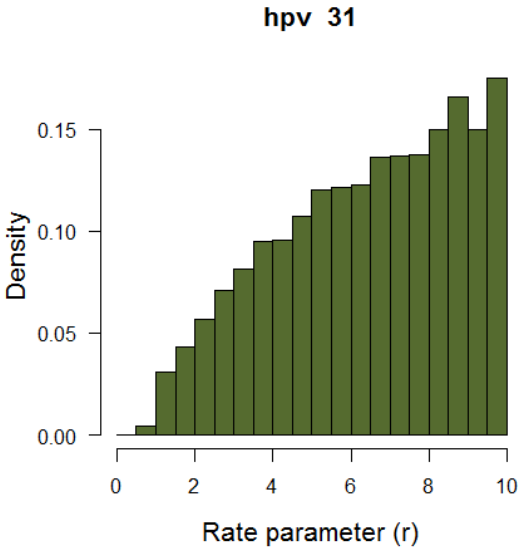


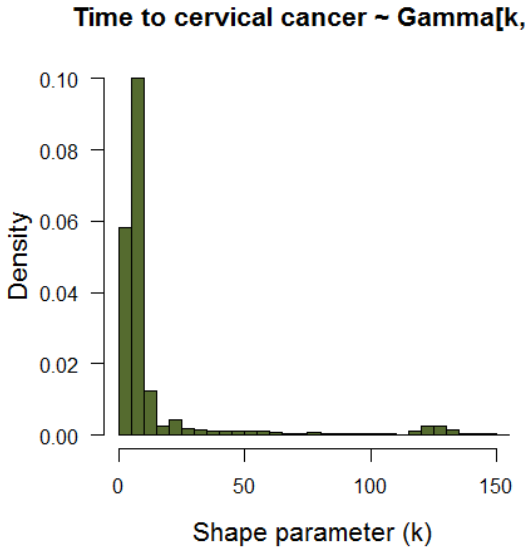
Figure A9 : Paramterisign progression to cancer using a gamma distribution, with increasing hazard with time. The shape is represented by parameter (p1), and the rate parameter is represented by p2. The subsequent plots show the posterior distribution derived for the waiting time to squamous cell carcinoma and adenocarcinoma in hpv type 16, 18,, 31, 33, 45, 51 and 52.

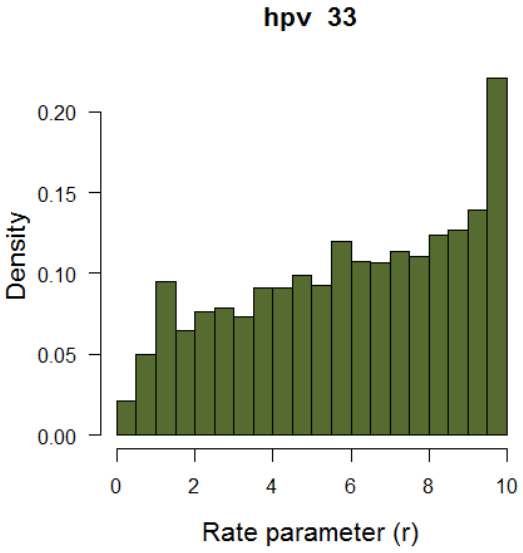


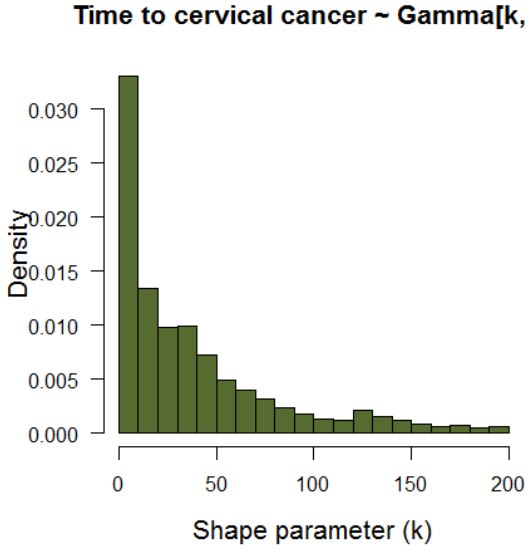


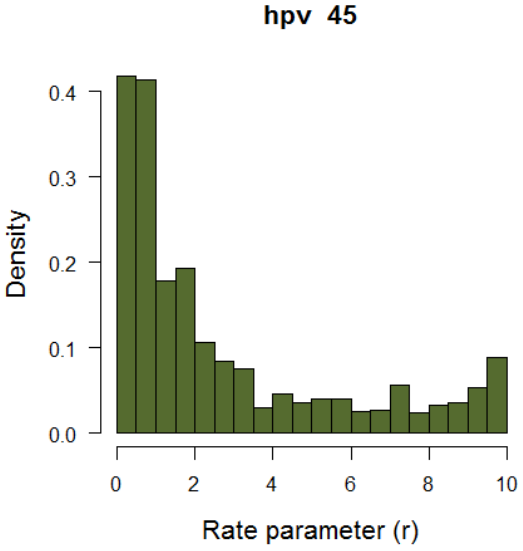


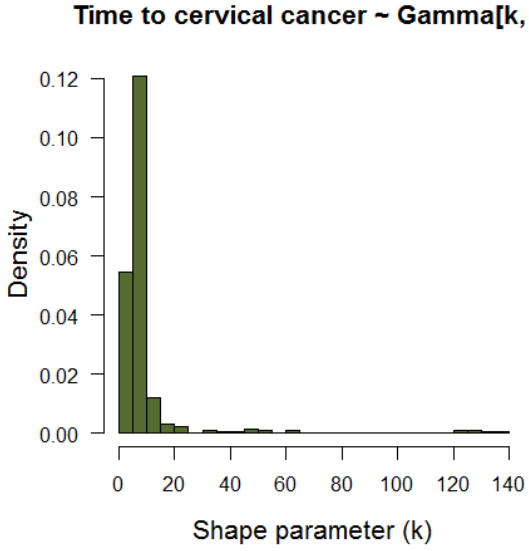


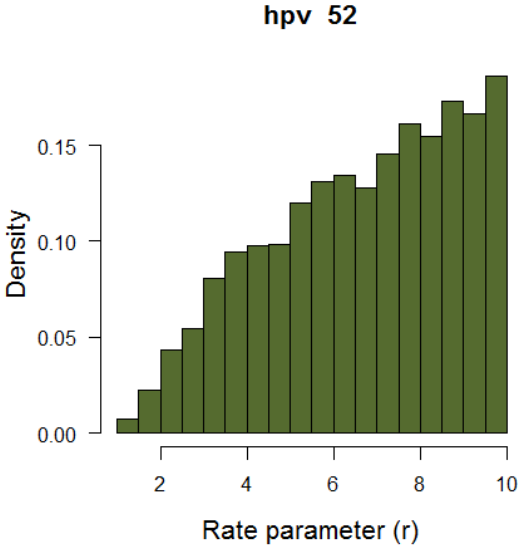

Time to cervical cancer ~ Gamma[k, r]
Density
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
Shape parameter (k)

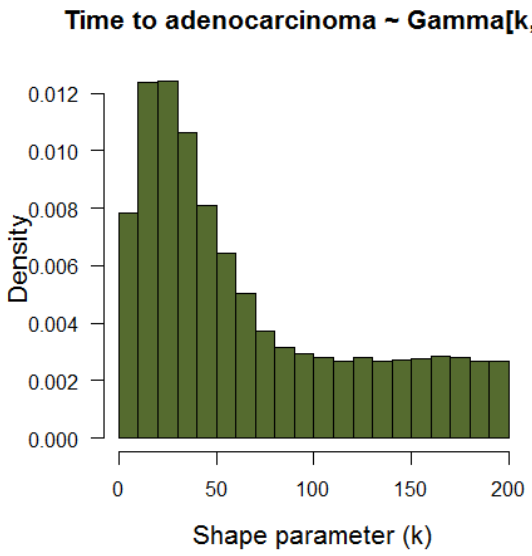

hpv 33
Density
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
Rate parameter (r)

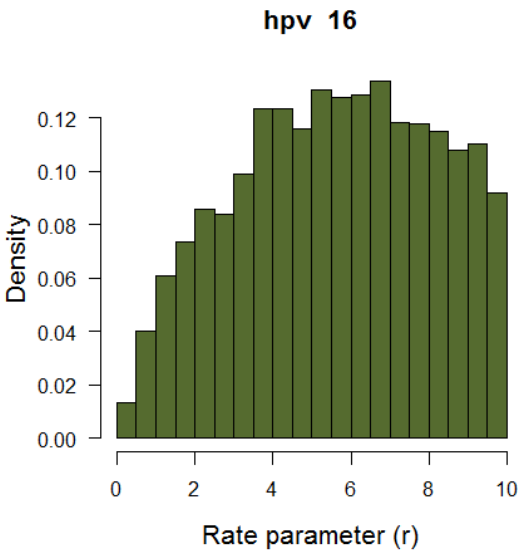

Time to cervical cancer ~ Gamma[k, r]
Density
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
50
100
150
200
Shape parameter (k)

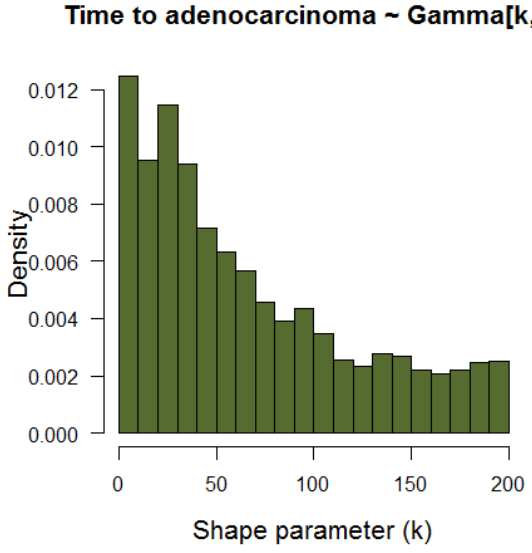

hpv 45
Density
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
10
Rate parameter (r)

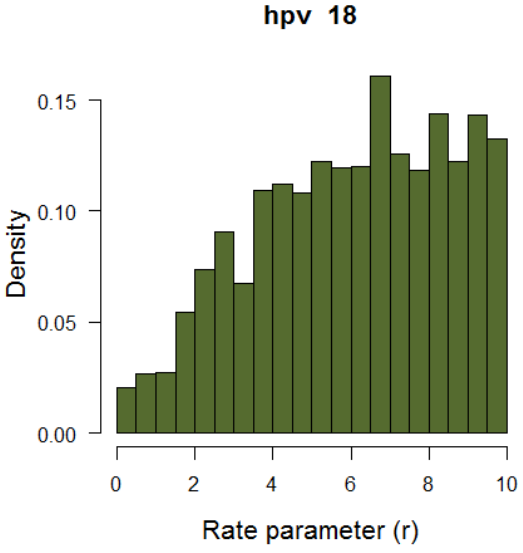

Time to cervical cancer ~ Gamma[k, r]
Density
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
20
40
60
80
100
120
140
Shape parameter (k)

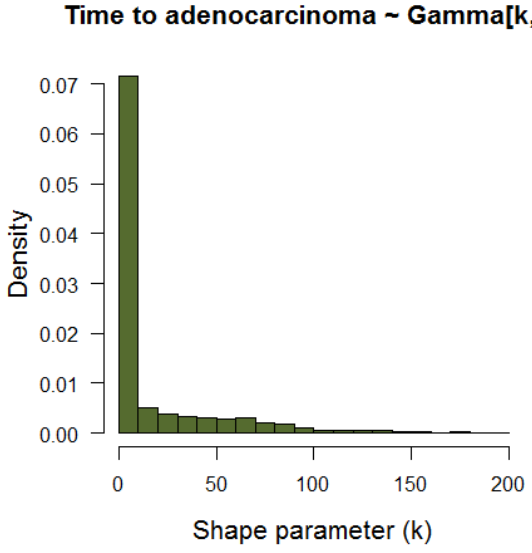

hpv 52
Density
0.15
0.10
0.05
0.00
2
4
6
8
10
Rate parameter (r)

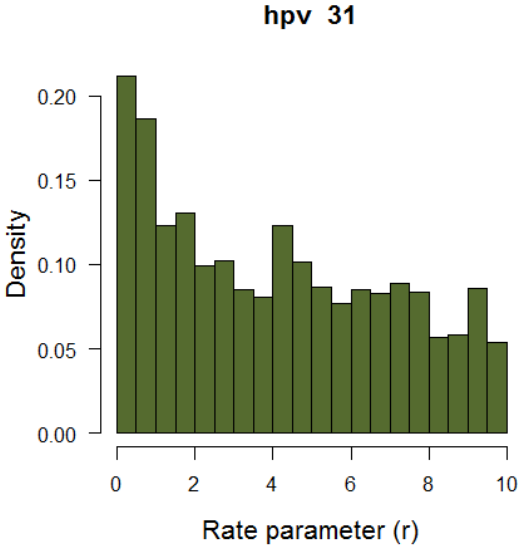

Time to adenocarcinoma ~ Gamma[k, r]
Density
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0
50
100
150
200
Shape parameter (k)

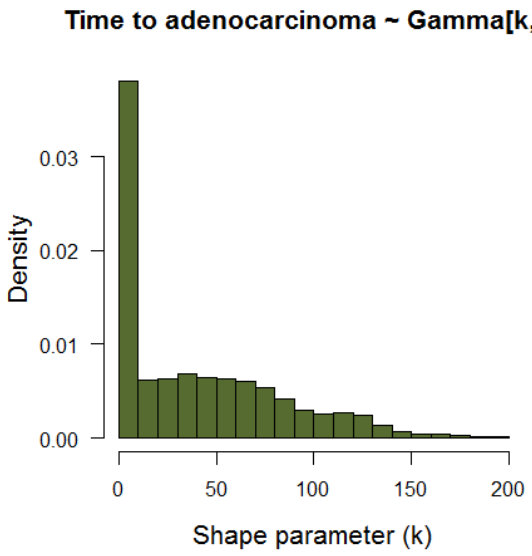

hpv 16
Density
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
2
4
6
8
10
Rate parameter (r)

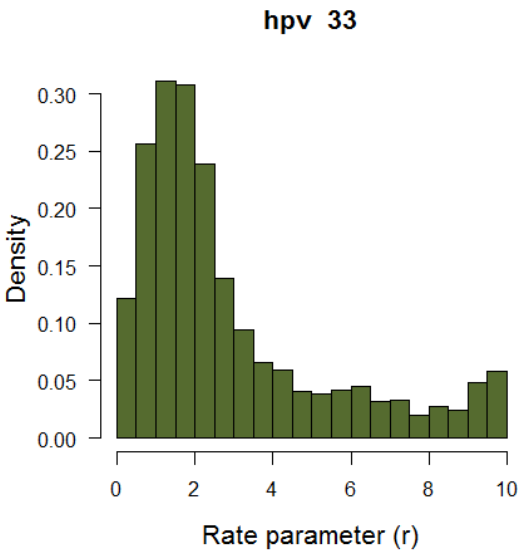

Time to adenocarcinoma ~ Gamma[k, r]
Density
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0
50
100
150
200
Shape parameter (k)

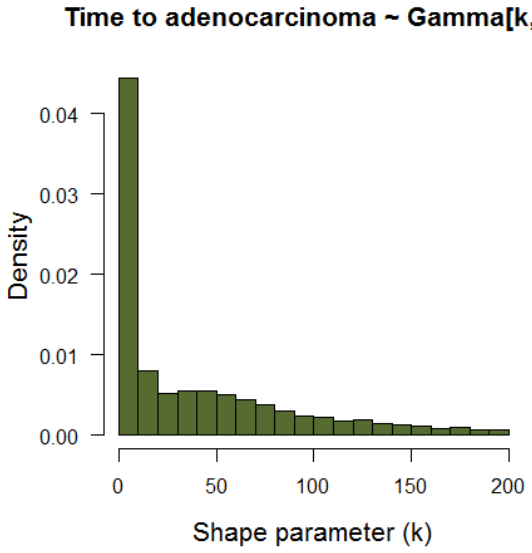

hpv 18
Density
0.15
0.10
0.05
0.00
0
2
4
6
8
10
Rate parameter (r)

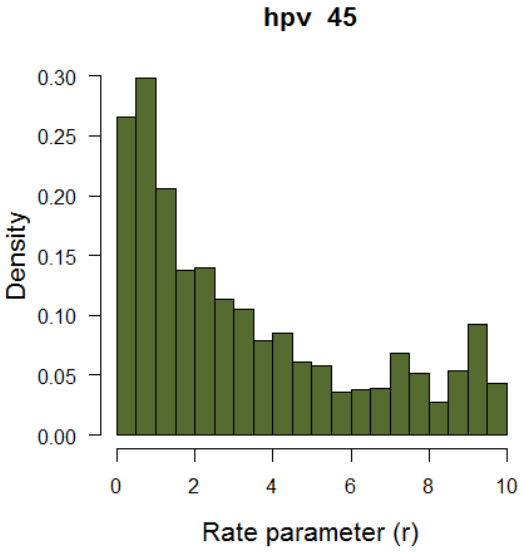

Time to adenocarcinoma ~ Gamma[k, r]
Density
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
50
100
150
200
Shape parameter (k)

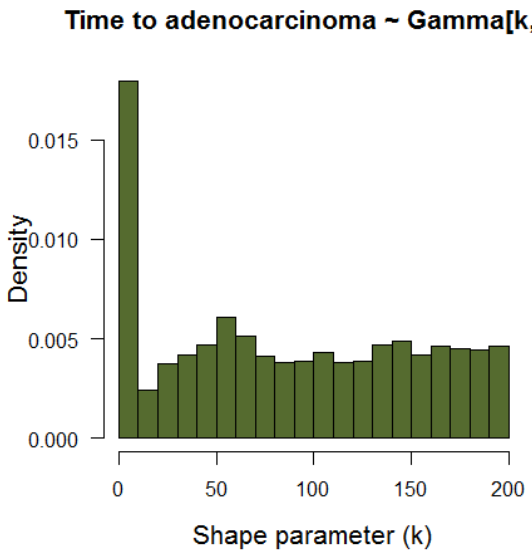

hpv 31
Density
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
Rate parameter (r)

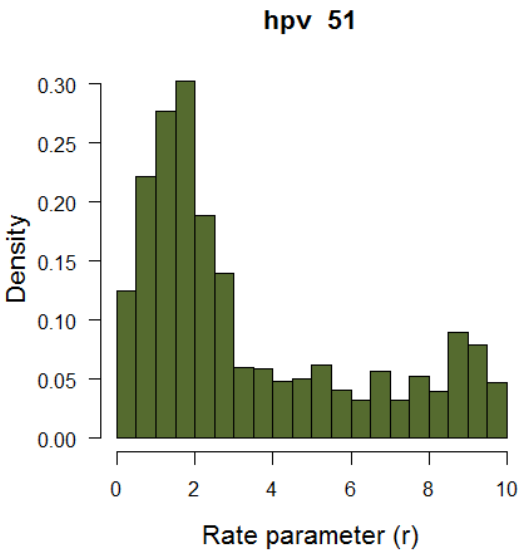

Time to adenocarcinoma ~ Gamma[k, r]
Density
0.00
0.01
0.02
0.03
0
50
100
150
200
Shape parameter (k)

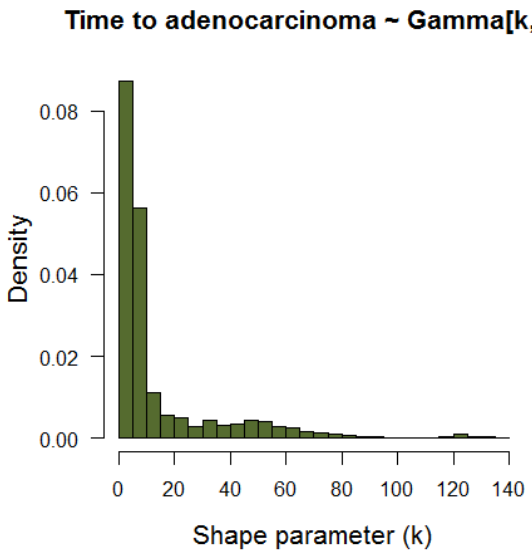

hpv 33
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
2
4
6
8
10
Rate parameter (r)

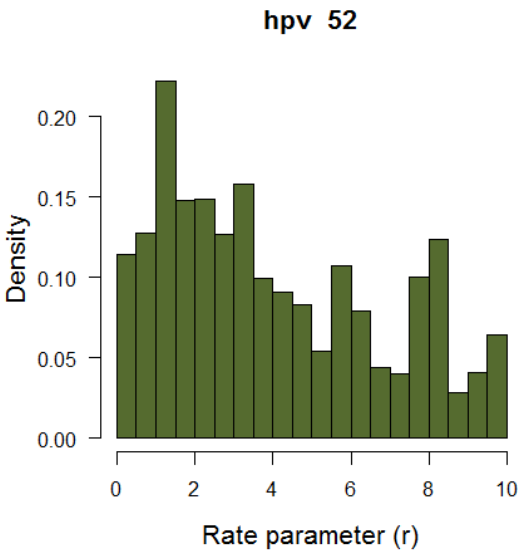

Time to adenocarcinoma ~ Gamma[k, r]
Density
0.00
0.01
0.02
0.03
0.04
0
50
100
150
200
Shape parameter (k)


hpv 45
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
2
4
6
8
10
Rate parameter (r)


Time to adenocarcinoma ~ Gamma[k, r]
Density
0.000
0.005
0.010
0.015
0
50
100
150
200
Shape parameter (k)


hpv 51
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
2
4
6
8
10
Rate parameter (r)


Time to adenocarcinoma ~ Gamma[k, r]
Density
0.00
0.02
0.04
0.06
0.08
0
20
40
60
80
100
120
140
Shape parameter (k)


hpv 52
Density
0.00
0.05
0.10
0.15
0.20
0
2
4
6
8
10
Rate parameter (r)

A4 Parameterisation of Screening Behaviour

Age at First Screen

A lognormally distribution is used to characterise the age at which a women attends her first cervical screen (for age 24.5 years and above) (Figure A10). A nelder-mead optimization is used to identify the distribution of age at first cervical screen that is best able to describe the observed fraction of women that have never attended screening with age [source: Cervical Screening Programme 2011-2012].

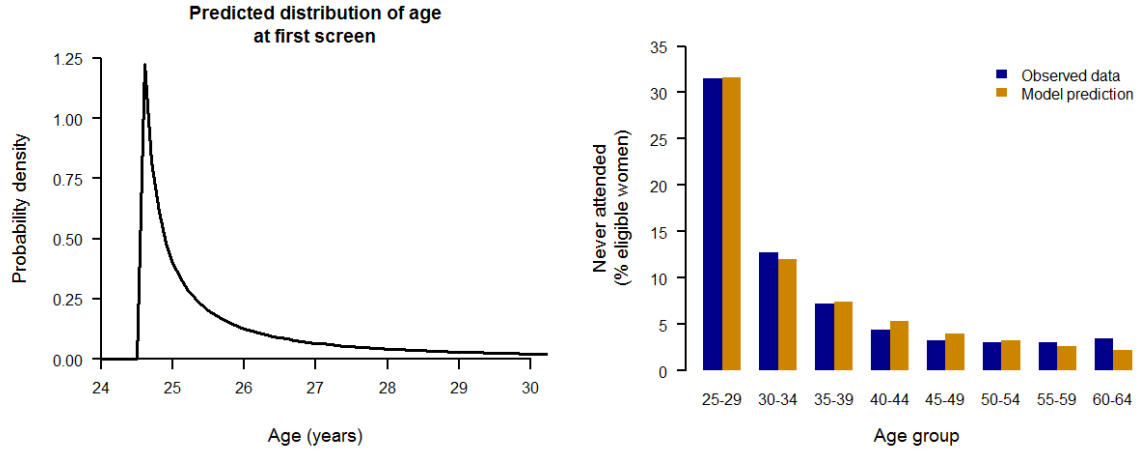
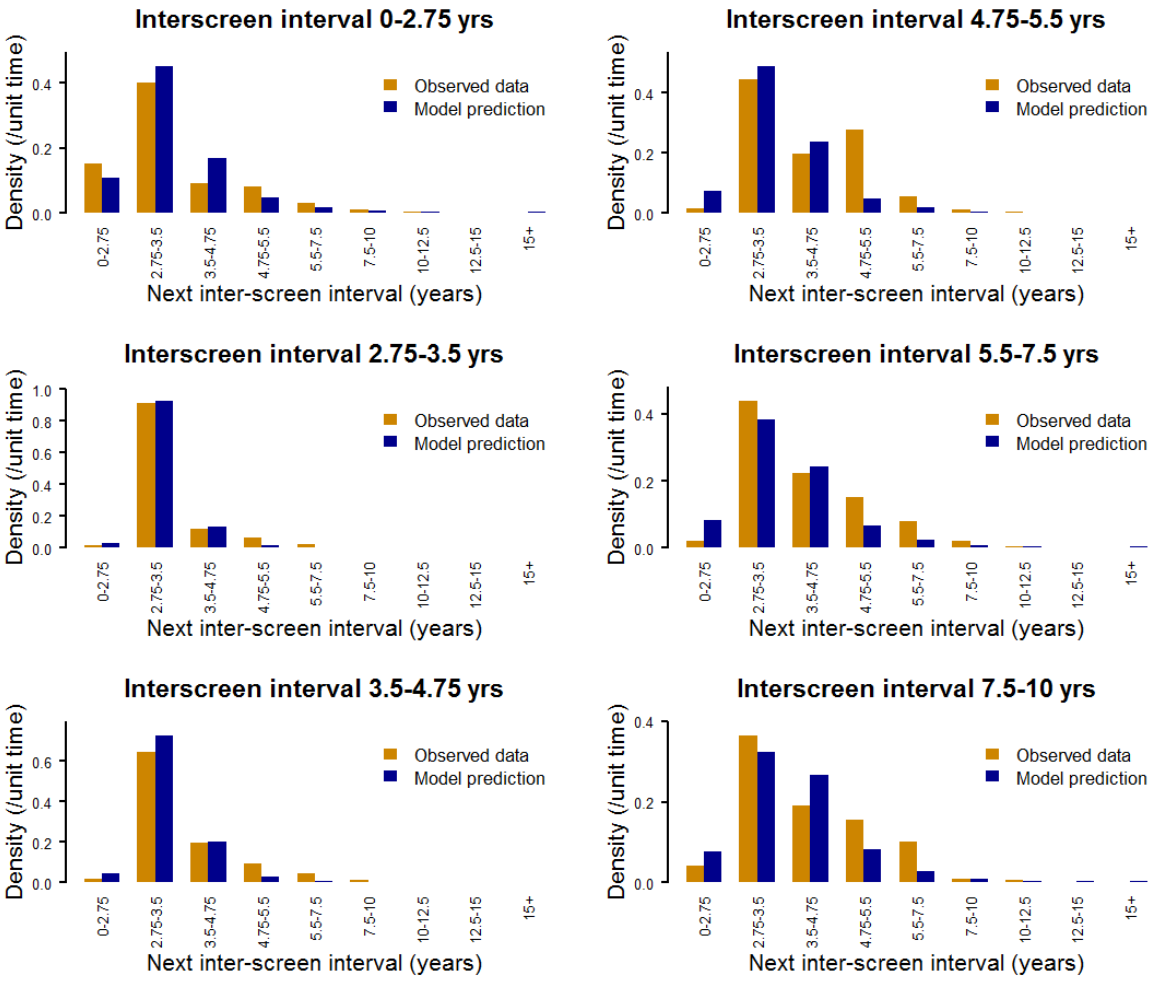


Figure A10: Age at first cervical screen**. Left panel**: Best-fit probability density (likelihood) of first attending screening at given age. **Right panel**: Model predictions and observed values for percentage of women that have never attended for screening, as a function of age.

Screening adherence

We use the time between two successive screens to identify long-term behavioural screening pattern. The data is restricted to women on routine recall with no history of abnormalities. A log-cauchy distribution of waiting time best describes the observed interval between successive screens. The interval between the last and penultimate screen is studied in women under 50 year with a prescribed interval of 3 years and data are stratified according to the previous inter-screening time (between screen (*n-2*) and *(n-1)*) – under 2.75, 2.75-3.5, 3.5-4.75, 4.75-5.5, 5.5-7.5, 7.5-10, 10-15 and 15 plus years and never screened). The observed distributions are fitted simultaneously to identify best-fitting log-cauchy scale and location parameters as a function of previous inter-screen. There is a slight increase in mean waiting time with increasing inter-screening time. This shift in expected value is accompanied by an accelerated increase in the variation of subsequent waiting times as individuals diverge form the 'prescribed' 3 year routine recall. In figure A11 we interpolate between the predicted mean and 90% interval to give a smooth distribution of inter-screening waiting that is then used to predict the time to next screen given an individual's screening history.



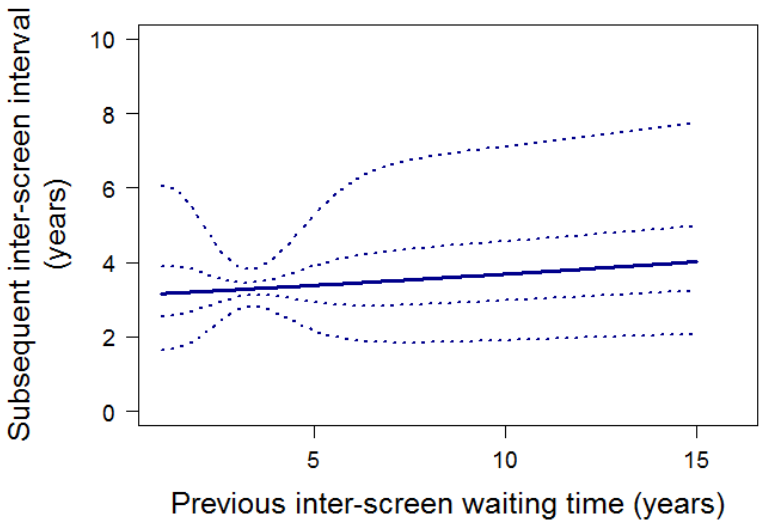


Figure A11: Distribution of waiting time to next screen. **Upper panel**: Observed and predicted distribution of 'next' inter-screen intervals when the previous interval was known to lie in the range 0-2.27, 2.75-3.5, 3.5-4.75, 4.75-5.5, 5.5-7.5 or 7.5-10 years. **Lower panel**: Best-fitting waiting time percentiles (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}) as a function of previous inter-screen interval.

A5 Economic costs

	Karnon (2003 HTA) Liquid based-cytology in cervical screening	Moss (2004) Evaluation of HPV/LBC Cervical Screening Pilot Studies	Kitchner (2011 HTA) MAVARIC - a comparison of automation-assisted and manual cervical screening	Legood (BMJ 2012) Cost -effectiveness of HPV in test of cure	Kitchner (2014 HTA) Clinical effectiveness and cost effectiveness of primary HPV screenign	Primary HPV pilot site 2014	NHS supplier chain recommended price 2014	Mean cost (sd)
Consultation / sampling	**10.25**	**13.63**	**16.35**	**18.99**	**17.31**			**15.31 (3.43)**
Cytology testing								
Consumables	5.17	5.09	3.96	7.78	5.45/15.4[i]	4.37		
Equipment and labour process slides	1.00	0.91						
Smear reading costs	2.57	2.26	1.37			13.63		
Other costs (overheads , staff time spent on admin, admin costs in lab, transport)	10.88	10.53	[redacted]	[redacted]	[redacted]			
Lab cost per cytology slide excluding other costs	8.74	8.27	5.33	7.78	6.445 / 8.44 [ii]			
Total lab cost per slide	**19.62**	**18.80**	***16.03****	***18.48****	***16.97/18.96**[#]*	**18**[iii]		**18.15 (1.3)**
HPV testing								
Consumables						6.5	6.31 6.35 6.19 7.49	
Equipment costs			17.22	13.88				
Staff costs - admininstration / sample preparation / performing test / quality control					4.5 / 9.38[iv]	3.25		
Other costs (overheads , staff time spent on admin, admin costs in lab, transport)			[redacted]	[redacted]			[redacted]	
Lab cost per sample excluding other costs			17.22	13.88				
Total lab cost per sample			***20.47****	***17.13****	**4.5 / 9.38**[iv]	**9.75**	***9.56** 9.6* 9.44* 10.74**	**9.75/12.01**[v] **(0.5/4.31)**

Table A4: Comparison of screening costs from historic economic analyses of the NHS cervical screening programme in England[7–11], inflated to 2014 values. *Missing costs (overheads /other lab costs) are estimated using average of available values. Notes: (i) costs quoted for normal / abnormal outcomes; (ii) costs for primary and triage calculated using expected proportion of abnormal outcomes; (iii) using a SurePath cytology assay, final value includes equipment, staff costs and all other lab overheads and costs; (iv) manufacturers estimated cost/ sentinel sites study cost; (v) costs using values from primary hpv testing studies only / including historical hpv triage test costs.

Treatment	**Cancer stage at diagnosis**					
	1A	**1B**	**2**	**3**	**4**	**1B+**
None	4.6%	5.4%	8.6%	12.3%	19.6%	19.8%
Cone	69.6%	18.1%	0.7%	1.0%	1.4%	15.8%
trachelectomy	1.0%	5.6%	0.2%	0.0%	0.0%	1.0%
hysterectomy only	20.4%	46.0%	7.6%	1.5%	1.4%	19.8%
radiotherapy (+/- hysterectomy)	1.5%	6.9%	20.7%	24.1%	32.4%	8.9%
chemotherapy (+/- hysterectomy)	0.4%	2.0%	3.8%	6.9%	10.1%	4.0%
chemo-radio therapy (+/- hysterectomy)	2.6%	16.1%	58.4%	54.2%	35.1%	30.7%

Table A5: Treatment of cancers according to stage at diagnosis (source: Cervical Cancer Audit, 2010)

Age at diagnosis	**Cancer stage at diagnosis**					
	1A	**1B**	**2**	**3**	**4**	**1B+**
25	47.8%	35.4%	6.2%	1.8%	3.5%	5.3%
25-49	48.9%	35.7%	7.7%	2.6%	1.3%	3.8%
50-64	21.3%	33.9%	17.9%	11.3%	7.3%	8.3%
65 and over	6.6%	27.7%	26.6%	17.9%	12.9%	8.4%

Table A6: Observed state of cancer progression, according to age at diagnosis (source: Cervical Cancer Audit, 2010)

References

1. Burington B, Hughes JP, Whittington WLH, et al. Estimating duration in partnership studies: issues, methods and examples. *Sex Transm Infect*. 2010;86:84-89. doi:10.1136/sti.2009.037960.

2. Howell-Jones R, Bailey a, Beddows S, et al. Multi-site study of HPV type-specific prevalence in women with cervical cancer, intraepithelial neoplasia and normal cytology, in England. *Br J Cancer*. 2010;103(2):209-216. doi:10.1038/sj.bjc.6605747.

3. Howell-Jones R, de Silva N, Akpan M, et al. Prevalence of human papillomavirus (HPV) infections in sexually active adolescents and young women in England, prior to widespread HPV immunisation. *Vaccine*. 2012;30(26):3867-3875. doi:10.1016/j.vaccine.2012.04.006.

4. Desai S, Chapman R, Jit M, et al. Prevalence of human papillomavirus antibodies in males and females in England. *Sex Transm Dis*. 2011;38(7):622-629. doi:10.1097/OLQ.0b013e31820bc880.

5. Scherpenisse M, Mollers M, Schepp RM, et al. Changes in Antibody Seroprevalence of Seven High-Risk HPV Types between Nationwide Surveillance Studies from 1995-96 and 2006-07 in The Netherlands. *PLoS One*. 2012;7(11). doi:10.1371/journal.pone.0048807.

6. Michael KM, Waterboer T, Sehr P, et al. Seroprevalence of 34 human papillomavirus types in the German general population. *PLoS Pathog*. 2008;4(6). doi:10.1371/journal.ppat.1000091.

7. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. *Health Technol Assess (Rockv)*. 2004;8(20):1-90.

8. Moss SM, Gray A, Marteau T, Legood R, Henstock E, Maissi E. Evaluation of HPV / LBC Cervical Screening Pilot Studies Report to the Department of Health (Revised October 2004). 2004;(October).

9. Kitchener HC R. B, H. C, et al. MAVARIC - A comparison of automation-assisted and manual cervical screening: A randomised controlled trial. *Health Technol Assess (Rockv)*. 2011;15(3):1-176.

10. Legood R, Smith M, Lew J-B, et al. Cost effectiveness of human papillomavirus test of cure after treatment for cervical intraepithelial neoplasia in England: Economic analysis from NHS Sentinel Sites Study. *BMJ*. 2012;345(7882):e7086.

11. Kitchener HC, Canfell K, Gilham C, Sargent A, Roberts C, Desai M. The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England. *Health Technol Assess (Rockv)*. 2014;18(23). doi:10.3310/hta18230.

YHEC

York Health Economics Consortium

Providing consultancy and research in health economics for the NHS, pharmaceutical and health care industries since 1986

UK National Screening Committee

Modelling studies addressing HPV screening

Draft Report v 1.1

JAMES MAHON, Senior Associate

STUART MEALING, Associate Director

19 January 2017

UNIVERSITY of York

Contents

Section 1: Introduction

1.1 BACKGROUND

Worldwide, cervical cancer is thought to be responsible for around 275,000 deaths per year which ranks it as the second most common cause of cancer deaths in women after breast cancer. Screening for pre-invasive changes in cervical epithelium, by means of cervical cytology, has resulted in a major fall in both incidence and deaths in the developed world. As a result of screening, cervical cancer now ranks 15th in female cancer deaths in the UK and mortality has fallen from 6.4 per 100,000 population in 1988 to 2.2 per 100,000 population in 2012.

Current policy recommendation

The UK NSC recently recommended that the National Cervical Screening Programme (NCSP) should change from a cytology strategy. The currently recommended strategy is based on detection of human papilloma virus (HPV) as the primary screening test followed by triage using liquid cytology in those with positive HPV tests. Women aged 25 – 64 are eligible for HPV based screening.

This primary change to the programme was based on trial evidence of clinical effectiveness, cost effectiveness evaluation and piloting in the UK.

The NCSP is now preparing to implement the new programme and policy recommendations are required on a number of key questions. These are identified below. While evidence is emerging on these questions direct evidence is very limited or non-existent. In countries which have adopted HPV based screening, policy development has therefore centred on outputs from modelling exercises.

Previous modelling of HPV in the UK

A UK model was developed by Bains in 2015. This unpublished study used disease transmission modelling to compare three strategies:

- Primary cytology followed by HPV triage (non-genotyped) of women with borderline/mild cytology results. Three-year and five year recall for negative women aged 25-50 and 50 years respectively.
- Primary high-risk (HR) HPV testing with cytology triage for HR HPV positive women with recall as above
- Primary HR HPV testing with cytology triage for HR HPV positive women with five year recall for negative women of all ages

Findings from the Bain's model suggested that compared to primary cytology with non-genotyped triage, either primary HR HPV testing strategy would be cost and life year saving. Five year recall for HR HPV would have the greatest saving (£35 million per year (£38 per woman) compared to £15.8 million (£14 per woman) with age dependent recall) but five year recall would generate a lower life year gain (0.0008 discounted life years compared to 0.0026 with age dependent recall).

Whilst the total saving per year is large, it is driven primarily by a lower cost of HR HPV testing compared to cytology. The discounted life year saved per woman with either HR HPV strategy were very small, at 0.3 days for five year intervals and 0.9 days with age

dependent intervals compared to primary cytology. Further, the impact on QALYs is not clear and was shown to be sensitive to how utility decrements were chosen and applied for screening and the results of screening. Whilst the Bain's model therefore seems robust in terms of HR HPV being cost and life year saving to the NHS in England compared to cytology with non-genotyped HPV triage, the modelling is unclear as to whether

- HR HPV is actually cost-effective in terms of the cost per QALY gained over a primary cytology strategy
- HR HPV is cheaper but less effective in terms of QALYs generated than a primary cytology strategy
- Fixed repeat screening intervals for all HR HPV negative women are cost effective compared to variable intervals by age

1.2 OBJECTIVE

Whilst the Bain's model provided information that could help guide the UK NSC on an appropriate screening strategy in the UK, the UK NSC wished to put the findings of the model in the context of other models published in the UK and internationally. In this regard the UK NSC wanted to explore three key areas:

i) Screening intervals for HPV negative women

The NCSP would like to recommend that primary HPV cervical screening should be offered:

- five yearly following a negative HR-HPV test for women aged 25-49
- ten yearly following a negative HR-HPV test for women aged 50-64 (or at 64 if aged 55-60 at the previous screening test).

As such, this study should identify what published models have reported on the clinical, cost and resource implications of these screening intervals in HPV negative women. The primary question is therefore:

- have five year screening intervals for HPV negative women been found cost effective in published models?

With the following sub questions:

- have published studies explored screening intervals by age group and if so what have they found?
- What is the duration of protection from cervical abnormalities in women over the age of 64 (or how was this been incorporated into models)?

ii) Surveillance recall intervals in HPV positive / cytology negative women

Diverging strategies have been proposed by the English and Scottish Screening Programmes for managing women in this group.

Both strategies would aim to recall women who screened HPV positive and cytology negative for HPV surveillance testing at 12 months.

- those who are HPV negative would return to screening at 5 yearly intervals
- those who are HPV positive and cytology positive would be referred to colposcopy

The strategies diverge on the use of HPV genotyping to inform the onward management of women with persistent HPV positive and cytology negative results:

- in Scotland, the proposal is that all women in this group would be recalled for repeat HPV testing in a further 12 months.
- in England, the proposal is that women with HPV 16 or 18 results would be referred for colposcopy. Women with all 'other type' HPV results would be recalled for repeat testing in a further 12 months.

An analysis of models addressing the clinical, cost and resource implications of these surveillance strategies is required to inform a UK NSC recommendation on this issue to answer the following questions:

- Have modelled estimates of surveillance recall strategies for women testing HPV + / cytology – identified an optimum approach?
- Have models compared approaches taking genotype (e.g. HPV 16 and 18) into account with those which do not?

iii) Options for women with HPV positive and cytology negative at the 'programme exit' test

The NCSP would also like to recommend that women who are HR-HPV positive at their final screening test should be recalled at 12 months and, if still HPV positive, be referred for colposcopy. If colposcopy is:

- decisively negative this would prompt discharge from the programme
- decisively positive this would prompt the offer of loop excision
- indecisive this would prompt the offer of loop excision or recall a further 12 months later.

An analysis of models addressing the clinical outcomes from this or other programme exit strategies is required to inform a UK NSC recommendation on this issue by answering the following questions:

- Have models explored programme exit strategies for women with positive HR-HPV results at the final screening test?
- What is the risk of cervical abnormalities developing in women who are HPV + / colposcopy-?
- should women who are HPV + / colposcopy indecisive be offered a choice of loop excision or further annual surveillance?
- how many rounds of annual surveillance should be offered to women who are HPV + / colposcopy indecisive?

Section 2: Methodology

A Rapid Evidence Assessment was undertaken to identify models published since 2005 that provided evidence against the key research questions.

2.1 PICOS

The PICOS for the evidence assessments for the three issues are provide in tables 2.1 to 2.3. In all cases studies were limited to full papers in peer reviewed journals (abstracts or posters were excluded), those in economically developed countries with a publication date of 2005 or later and to English language studies only. Only studies considering a screening age starting at 25 or older were included.

Table 2.1: Issue 1 (Screening intervals for HPV negative women) PICOS

Question	Have modelled estimates found five-year routine screening intervals for HPV negative women to be effective?
Sub-questions	Has interval variation by age group been modelled? What is the duration of protection against cervical abnormalities in women older than 64 years?
Population	Women screened for cervical cancer
Intervention	HPV based screening 5-year screening intervals
Comparator	Cytology based screening HPV based screening with different interval duration
Outcomes	Modelled outcomes including the following where reported: <u>Clinical measures</u> Cumulative incidence of CIN2+, CIN3+ and cancer Mortality Treatment of precancerous lesions / cancer prevented <u>Service resource use, expected number of:</u> Cytology tests HPV tests Colposcopies Histology evaluations Treatment for precancerous lesions Treatment of cancer <u>Lifetime / individual expected number of:</u> Screening / follow up episodes Colposcopies <u>Cost effectiveness, cost comparisons between strategies</u>
Study types	Economic evaluations with modelling

Table 2.2: Issue 2 (Surveillance recall intervals in HPV positive / cytology negative women) PICOS

Question	Have modelled estimates of surveillance recall strategies for women testing HPV + / cytology – identified an optimum approach?
Sub-questions	Have models compared approaches taking genotype (e.g. HPV 16 and 18) into account with those which do not?
Population	Women with HPV + / cytology – screening test results
Intervention	HPV based screening 'untyped'
Comparator	Cytology based screening HPV based screening using different strategies e.g. genotyping
Outcomes	Modelled outcomes including the following where reported: <u>Clinical measures</u> Cumulative incidence of CIN2+, CIN3+ and cancer Mortality Treatment of precancerous lesions / cancer prevented <u>Service resource use, expected number of:</u> Cytology tests HPV tests Colposcopies Histology evaluations Treatment for precancerous lesions Treatment of cancer <u>Lifetime / individual expected number of:</u> Screening / follow up episodes Colposcopies <u>Cost effectiveness, cost comparisons between strategies</u>
Study types	Economic evaluations with modelling

Table 2.3: Issue 3 (Options for women with HPV positive and cytology negative at the 'programme exit' test) PICOS

Question	Have models explored programme exit strategies for women with positive HR-HPV results at the final screening test?
Sub-questions	i) what is the risk of cervical abnormalities developing in women who are HPV + / colposcopy –? ii) should women who are HPV + / colposcopy indecisive be offered a choice of loop excision or further annual surveillance? iii) how many rounds of annual surveillance should be offered to women who are HPV + / colposcopy indecisive?
Population	Women 64 years of age with HPV infection in post screening surveillance rounds.
Intervention	i) discharge from routine HPV based screening in women who are HPV + / colposcopy – ii) loop excision or annual surveillance in women who are HPV + / colposcopy indecisive iii) multiple rounds of annual surveillance of women who are HPV + / colposcopy indecisive
Comparator	Any other strategy for women aged 64 years of age HPV+ in post screening surveillance rounds
Outcomes	Incidence of abnormalities requiring management in the above groups. Abnormalities are: • CIN2+ • CIN3+ • Invasive cervical cancer
Study types	Economic evaluations with modelling

2.2 LITERATURE SEARCH

2.2.1 Search Strategy

A search strategy was developed in conjunction with PHE to identify suitable studies (Appendix A). The search was conducted in Medline, Embase and the Cochrane Library on 20 November 2017.

2.3 STUDY SELECTION

As a REA study selection was undertaken by one reviewer.

2.4 DATA EXTRACTION

The following items were extracted from each study:

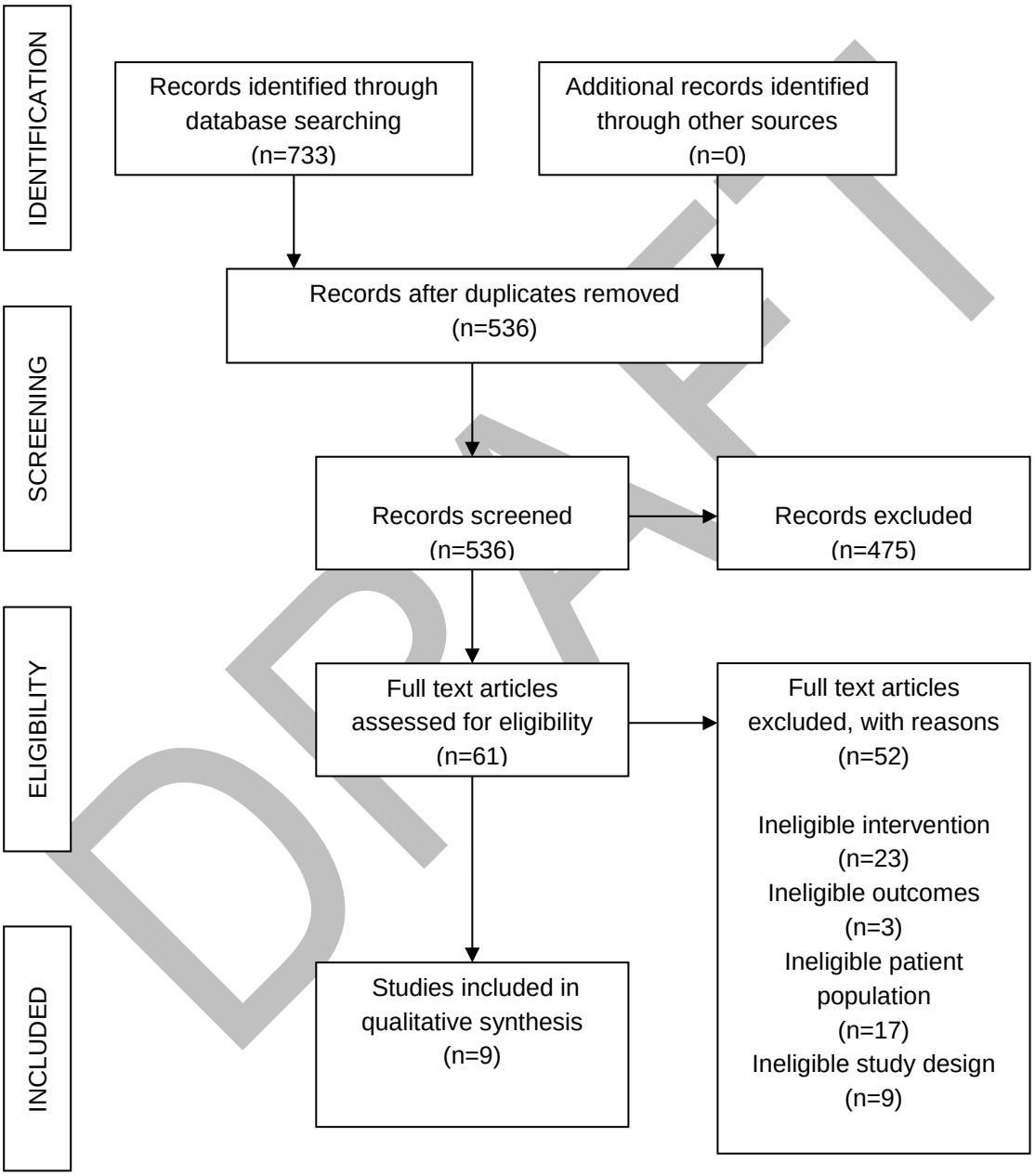
- Population in which the modelling took place;
- Country and setting;
- Detailed description of screening strategies compared (including frequencies);
- Description of modelling approach

    - Time horizon
    - Type of model
- Study results (including outcomes as set out in tables 2.1 to 2.3

3.1 STUDIES IDENTIFIED AND SELECTED

In total 733 potential records were identified across the three databases with 536 unique records. The record selection process is shown in Figure 3.1.

Figure 3.1: Study selection process



Section 4: Results

4.1 SUMMARY OF IDENTIFIED STUDIES

Nine studies were identified that met the inclusion criteria for the three issues to be addressed by the REA. Of these studies, one was in England[5], two in the Netherlands[2,4] and one each in Australia[8], Canada[6], Germany[9], Italy[1], New Zealand[7] and Norway[3].

All published studies essentially used the same broad modelling structure with some form of microsimulation of disease transmission coupled with markov processes of disease progression and decision analytical process for the outcomes from screening with calibration of model findings to published national epidemiological data. All studies had to use assumptions on compliance with screening for HPV acknowledging that this is unknown and may vary by the risk of HPV infection.

All but one study[9] were lifetime models with two studies[3,5] explicitly pointing out that findings from HPV screening studies from one country would not be generalizable to another due to differences between countries in the natural history of HPV infections (including age related incidence and transmission rates) and screening adherence. If this is the case the one study of an English population[5], provides results with the most relevance to the research issues.

A summary extraction table with key methods and findings is presented in Table 4.1 with full extraction tables in Appendix B.

4.2 ISSUE 1: SCREENING INTERVALS FOR HPV NEGATIVE WOMEN

All nine studies provided evidence for Issue 1, with the cost effectiveness of 5-year HPV screening compared to cytology assessed in all studies and the cost effectiveness of different intervals of HPV screening assessed in eight studies[1-6,8,9].

All studies – including the one English study - concluded that five-year HPV screening was an efficient strategy compared to cytology screening, with HPV screening costing less and in most scenarios considered having better outcomes than cytology. There was some evidence that in some scenarios outcomes were not always better with HPV screening with QALY losses in three models[5,7,8] . These scenarios were where a significant utility decrement for abnormal readings was applied. Whilst five-year HPV screening was considered to be efficient in all studies, two[1,6] studies looking at three-year intervals suggested that three-year screening was the optimal strategy. The difference in lifetime QALYs in all studies (when reported) between different screening intervals was around 0.005 per woman

In unvaccinated women, the English study[5] suggests that, depending on the follow up strategy for HR HPV positive women, five and six-year HR HPV (any oncogenic infection) screening strategies would result in a reduction in lifetime costs compared to cytology of £13-£16 and £24-£27 respectively. The model suggest QALY gains of 0.0003 to 0.0008 with five-year screening and 0.0007 to 0.0013 with six-year screening. The cost differential and QALY gains with five-year screening reported are therefore broadly in line with those from the Bain's model.

Six-year screening to age 49 followed by ten-year screening to age 64 would be £4 less costly than six-year screening for all ages with a QALY gain of 0.0004. This suggests that six-year screening to 49 followed by ten-year screening to age 64 is likely to be cost and QALY saving compared to six-year screening although the QALY differentials are very small between strategies. It is noted that a six-year HR HPV screening strategy may result in a small life year loss for six year followed by ten-year screening at age 49 both resulting in slight losses in life years (0.0004 to 0.0005 LY) compared to cytology screening.

For HR HPV screening using partial 16/18 genotyping in unvaccinated women, the English study provided evidence that five and six-year screening intervals would result in a reduction in lifetime costs compared to cytology of £3 and £15 per women respectively. HR HPV screening with partial genotyping would result in a QALY gains of 0.0009 with five-year screening and 0.0020 with six-year screening depending on the follow up strategy for HR HPV positive women. For six-year HR HPV screening in unvaccinated women to age 49 followed by ten-year screening to age 64 estimated that such screening intervals would be £4 less costly than six-year screening for all ages with a QALY gains of 0.0022. Six-year screening to age 49 followed by ten-year screening to age 64 would seem to be cost effective compared to six-year screening for all ages. Again, however, it is noted that the QALY gains are very small and could result in a very small loss in life years (0.0001) for a HR HPV with partial genotyping screening strategy compared to cytology.

In summary, the English study results would seem to provide evidence that for unvaccinated women HR HPV screening every six years until age 49 with ten-year screening to age 64 would be the most cost-effective strategy, but the result is somewhat ambiguous due to the very small QALY differences between all strategies and the loss in life years with this strategy (albeit again small) compared to cytology screening.

For vaccinated women, the English study reported almost identical results regardless of whether HPV non-genotyping or HR HPV testing was undertaken. Six-year screening would result in a QALY gain of 0.0021 to 0.0022 QALYs compared to cytology screening with a cost saving per woman of £28 to £29. Five-year screening would have a lower QALY gain of between 0.0001 and 0.0002 QALYs with a lower cost saving per woman of £15 to £16. Six-year screening (HPV non-genotyping or HR HPV testing) to age 49 followed by ten-year screening to age 64 would result in a QALY gain compared to six-year screening (HPV non-genotyping or HR HPV testing) of all ages of 0.0004 and a cost saving of £4 to £5.

As was the case for unvaccinated women, six-year HPV screening strategy to age 49 followed by ten-year HPV screening to age 64 was the most cost effective strategy for

vaccinated women although the use of HPV genotyping or not is essentially immaterial to the cost effectiveness results. However, it is noted again that the results are driven by very small differences in QALYs between strategies and that such an age dependent strategy always resulted in a slight loss in life years of 0.0003 per woman compared to cytology screening despite generating the highest QALY gain.

4.3 ISSUE 2: SURVEILLANCE RECALL INTERVALS IN HPV POSITIVE / CYTOLOGY NEGATIVE WOMEN

Four studies[3,6,8,5] assessed the cost effectiveness of different strategies for women who are HPV+/cyt-.

Three of these studies[5,7,8] assessed genotyping and the two which were outside the UK[7,8] concluded that management of HPV+ women using genotyping was the most cost effective approach.

The English study[5] assessed 12 month recall for HPV+/cyt- women against 24 month recall with differing strategies for HPV+ women by genotype.

For unvaccinated women, the shorter recall period was found to cost in the region of an additional £15 to £17 over a woman's lifetime compared to the longer recall with a QALY loss of between 0.0012 and 0.0016 with a 12 month as opposed to 24-month recall interval. However, the shorter recall did result in a life year gain of between 0.007 and 0.0011 depending on the genotyping strategy.

For vaccinated women, the results were essentially identical for unvaccinated women with shorter recall intervals incurring additional costs of £11 regardless of genotyping strategy with QALY losses of between 0.0013 and 0.0014 but life year gains of 0.0003 to 0.0004 compared to the longer recall intervals.

The evidence on 12-month recall compared to 24 month recall for HPV+/cyt- women from the English study would suggest that the shorter interval is likely to be costlier than the longer interval, but in terms of effectiveness the evidence is somewhat contradictory with a loss in QALYs but a gain in life years with 12 compared to 24-month recall. These results are independent of whether and how HPV genotyping is used in the overall screening strategy.

The findings in England on shorter recall periods are supported by the findings from a Norwegian study[3] that reported that a 6, 12 or 18-month recall period for HPV+ women make quite significant differences to cancers prevented and treated although minimal difference to the lifetime cervical cancer rate. The study did not report the discounted costs of different strategies and so did not report on the relative cost effectiveness of different recall intervals.

4.4 ISSUE 3: OPTIONS FOR WOMEN WITH HPV POSITIVE AND CYTOLOGY NEGATIVE AT THE 'PROGRAMME EXIT' TEST

No studies were identified that provided evidence for differential strategies for women on exit.

Table 4.1 Summary data extraction table

Paper	Country	Model structure	Screening strategies considered	Screening schedules (years)	Summary of cost effectiveness results	Study recommendation
Accetta 2010[1]	Italy	Markov with nine health states and death: Healthy, HPV infection (lrHPV, hr HPV 16/18, hrHPV non-16/18) pre-cancer lesions (low and high severity) and cancer (local, regional, distant). Women progress through model one at a time with annual cycles with state dependent probabilities	No screening, Cytology, primary HPV, Cytology followed by HPV triage, HPV followed by cytology triage	3 and 5	Current strategy of primary cytology every three years is dominated by primary HPV with cytology triage every three years. Five-year screening would be less expensive but with slightly worse outcomes.	Three-year screening with HPV with cytology triage
Berkhof 2010[2]	Netherlands	Markov with six monthly cycles. Health states not well described but can be inferred to include HPV free and HPV states with HPV low risk and high-risk states. Patients can develop CIN2+ only if in the HPV high risk state but can develop CIN1 in all HPV states. Progression to CIN3 was age dependent. Patients can progress from CIN states to cancer or back to a well state.	Cytology, HPV with cytology triage, Combination cytology and HPV, Cytology with HPV triage.	5, 6, 7.5 and 10	Strategies with a screening interval over 7.5 years were not cost effective with a willingness to pay threshold of €20,000/QALY. The optimal strategy was five-year screening with HPV followed by cytology triage	Five-year screening with HPV with cytology triage
Burger 2017[3]	Norway	Model starting at age 8 where girls/women have a probability over time of type-specific HPV incidence and clearance. This can progress to lesions and cancer which are a function of age, lesion and duration of infection. The model is stratified by HPV genotype, CIN grades and cancer stage.	**1. HPV followed by cytology triage** HPV+Cyt- women are re-tested after 12 months. HPV+ women at this point have colposcopy. HPV- women return to previous screening period. Wait time between re-testing between 6 and 18 months was explored as was 1,2 or 3 HPV+cyt- results required before referral for	3, 4, 5, 6, 8 and 10	Lengthening the time between screenings did have impact on the cancer incidence rate with more frequent screening reducing the cancer rate. However, the most important factor was starting screening at age 25. Different intervals between re-testing of HPV+/cyt- women or altering the point at which	HPV-based screening among unvaccinated women should start at age 25 with an appropriate use of cytology triage to control colposcopy referrals. No recommendation was made on the frequency of testing

Paper	Country	Model structure	Screening strategies considered	Screening schedules (years)	Summary of cost effectiveness results	Study recommendation
			colposcopy. **2. Cytology with HPV triage (current strategy)** **Cyt+(minor lesions) and** HPV+ (for high risk HPV) women are re-tested with cytology and HPV test after 6-12 months. High grade lesions on cytology are referred for immediate colposcopy.		colposcopy occurs for these women made little impact on the cancer incidence rate	
de Kok 2012[4]	Netherlands	Patient simulation model where women have a probability over time of type-specific HPV incidence and clearance. This can progress to lesions (which can clear) and cancer which are a function of age, lesion and duration of infection. The model is stratified by high risk HPV (but not specific genotype), CIN grades and cancer stage.	Nine strategies were considered with 171 policy combinations of start age and screening frequency. The only strategies considered with results reported were cytology, HPV with cytology triage followed by a second cytology triage at 6 months for HPV+/cyt- women with HPV sensitivity of 90% and 95%	3 to 10. Only results for 5 years presented	The cost effectiveness results of strategies were not presented. The summary of the results states in most scenarios primary HPV screening is the preferred scenario in women over 30	Where screening is well controlled, European countries should switch from cytology to HPV screening
Kulasingham 2009[6]	Canada	Poorly described but a cohort model with yearly cycles where women can move from healthy to precancer, cancer and death. Women can move from diseased back to healthy states	Eight strategies considered for a start age of 25 (a further 19 strategies for people under 25). HPV testing only, cotesting, cytology with HPV triage, HPV with cytology triage.	1, 2, 3 and 5	HPV testing every three years followed by cytology triage may be more effective and less costly than cytology screening alone	HPV with cytology triage from age 25 with three-year screening
Petry 2017[9]	Germany	Model is not described beyond being a cohort model with a decision tree component. No detail of health states was provided.	Multiple HPV strategies considered in terms of type of HPV test and cytology and also cotesting. Only results for cytology (annual) and HPV with cytology triage (3 and 5 years) are extracted	1 (cytology), 3 and 5 (HPV with cytology triage)	Screening strategies for HPV results in fewer cancers at a lower cost than cytology alone. Screening of HPV at intervals less than five years does result in more	No recommendation for a specific HPV screening strategy

Paper	Country	Model structure	Screening strategies considered	Screening schedules (years)	Summary of cost effectiveness results	Study recommendation
					cancers detected but at a higher cost than five-year screening	
Lew 2017[8]	Australia	Women cycle through following states: susceptible, HPV infected (and genotype), CIN1-3 and cancer (6 stages). Women can also be vaccinated or become immune. At each screening point the screening and treatment model is applied with a probability of screening attendance. Modules within the screening and treatment model include a colposcopy, biopsy and treatment modules and post treatment natural history module. Model outcomes were calibrated to observed data on cancer, cancer death rates and histology and abnormality rates	Seven strategies considered. Only extracted were cytology (with 5 years screening), HPV with cytology triage or HPV genotyping. All HPV+ women at discharge from screening are offered colposcopy. Current cytology practice is 3 years cytology screening (5 years at 45) Various other elements of screening strategies were considered including: different options for HPV+/low grade cytology (direct colposcopy or reflex HPV); whether women are invited to attend a first screening or not; different levels of compliance with call and recall programmes	2 (cytology), 5 and 6 (HPV)	The authors conclude that HPV testing every five years with partial genotyping or cotesting with cytology were the most effective. Sending those with HPV16/18 for colposcopy and other genotypes for reflex cytology was described as "one of the most cost-effective" strategies. Whilst the analysis is unambiguous that all strategies will result in lower cost and HPV strategies are likely to dominate non-HPV strategies (at least if only life years and not QALYs are considered) there is no full incremental analysis of strategies, QALY gains are small across strategies and may be negative for some HPV strategies and a wide range of different scenarios were undertaken making it difficult to isolate the actual effect of different aspects of strategies.	HPV testing every five years with partial genotyping and direct colposcopy if 16/18
Kitchener 2014[5]	England	Same model as Lew 2017	HPV with cytology triage and cytology alone are two main strategies. Within the HPV triage there are sub strategies	5 and 6 years and 6 years 25-49 followed by 10 years 50-64	HPV testing is a cost-effective strategy compared to cytology. Whilst most of the	The most feasible and cost-effective strategy in terms of delivery could involve a single policy

Paper	Country	Model structure	Screening strategies considered	Screening schedules (years)	Summary of cost effectiveness results	Study recommendation
			depending on the treatment pathway should a woman screen as HPV+/cyt negative. **Strategy 1**: HPV+/cyt- women are recalled for HPV with cytology triage in 24 months. HPV+/cyt- women are again recalled at 24 months **Strategy 2**: Initial screen is for HPV genotype. HR HPV+/cyt- women are recalled for HPV genotype with cytology triage in 24 months. 16/18 positive women are referred to colposcopy. Other HR+ (OHR) are referred for cytology with cyt- women again recalled at 24 months. **Strategy 3**: Initial screen is for HPV genotype. 16/18 positive women are referred to colposcopy. HPV+/cyt- women are recalled for HPV genotype with cytology triage in 24 months. Other HR (OHR)+ are referred for cytology with cyt- women again recalled at 24 months and move onto a 24-month retest cycle whilst they remain OHR+/cyt-. 16/18 positive women at retest are referred to colposcopy In all strategies, recall at 12 rather than 24 months was considered.		strategies considered were cost and QALY saving, they all resulted in greater numbers of colposcopies and biopsies in unvaccinated women. The QALY gains per woman were small with any strategy although primary HPV genotype testing only appears to be an efficient strategy in vaccinated women.	across the screening age range with 5- or 6-yearly screening intervals and 12-month recall for HPV positive women with negative cytology.

Paper	Country	Model structure	Screening strategies considered	Screening schedules (years)	Summary of cost effectiveness results	Study recommendation
Lew 2016[7]	New Zealand	Same model as Lew 2017	"Cytology, HPV with cytology triage, HPV with genotyping. Within the HPV triage there are sub strategies depending on the treatment pathway should a woman screen as HPV+/cyt negative. Strategies for HPV+ women are as follows **HPV with cytology triage**: HPV+/cyt- women are recalled for HPV and cytology cotest in 12 months. HPV+ or cyt+ women are sent for colposcopy. HPV-/cyt- back onto normal screening cycle **HPV with genotyping**: HPV16/18 sent for colposcopy. OHR+ are referred for cytology with cyt- women recalled at 12 months. HPV+ women at recall are referred to colposcopy and HPV- women back onto normal screening cycle	3 (cytology), 5 (HPV)	At a WTP threshold of $50,000/LY, in both unvaccinated and vaccinated women HPV genotyping was the most cost-effective strategy. When QALYs were considered (although detailed findings not presented in body of report) findings are reported to vary widely. If disutility for screening and/or a minor disutility for abnormal findings are considered, then HPV genotyping remains the cost-effective choice. If there is no disutility from screening itself but a major disutility from abnormal findings then all HPV strategies are less effective than cytology screening	Primary HPV with genotyping

Section 5: Discussion

This rapid evidence assessment identified nine studies that provided evidence against two of the three key issues it was designed to address. Evidence is available on the cost-effectiveness of five-year HPV screening and on differential approaches to strategies for HPV+/cyt- women but no studies reported differential cost effectiveness results for different strategies for women on exit from screening.

Whilst not a key research question, where it has been looked at by published studies HPV vaccination does not seem to significantly influence the relative cost-effectiveness of HPV versus cytology screening (i.e. if HPV was found to be efficient with a strategy of no vaccination it was also found to be efficient if vaccination was being undertaken).

Two key findings or conclusions can be drawn in relation to the original research questions.

Key finding one: Five-year HPV screening is reported as being an efficient strategy compared to cytology screening but the true cost-effectiveness of this (and the optimal screening period) is uncertain

All published models have reported that HPV screening is likely to be an efficient strategy compared to cytology screening. This is in line with the previous unpublished UK model.

The finding that HPV screening is efficient compared to cytology screening is driven by the higher sensitivity/specificity of HPV compared to cytology and the lower number of screenings that are required with HPV screening. This, in turn, is the main driver towards the conclusion that HPV screening will be cost saving. However, the impact on outcomes is more ambiguous than may be suggested by the economic models both published and unpublished for the following reasons:

- Life year gains were small in all studies as were the absolute number of cancers reduced as a percentage of the total population screened. For example, one study, found that more frequent cytology screenings may reduce mortality compared to less frequent HPV screening. Similarly, depending on utility values chosen for the results of an abnormal screening, some studies found that HPV screening could result in a reduction in QALYs compared to cytology screening.

  As stated explicitly by the one study from England[5] findings from studies of HPV screening strategies in one country are unlikely to be transferable to another given the differences between countries in HPV prevalence, the natural history of HPV in populations based upon lifestyle choices and the adherence by women to screening intervals. This means that in studies other than the English study, it is likely that the costs and outcomes reported for different strategies have limited generalisability to the UK context.

- Costs and benefits of HPV screening are driven by assumptions that have to be made on the compliance rates for screening intervals for what is essentially an STD. One Dutch study[4] found that for HPV screening to be cost-effective HPV screening had to be 'well controlled'. How this will differ for all women and for those who see themselves as low and high risk is currently unknown.

- The numbers of colposcopies could go up or down with HPV screening depending on the exact nature of the screening strategy and model assumptions employed. The exact direction (an increase or decrease in colposcopies) was uncertain with the unpublished UK model predicting a rise in colposcopies with HPV screening with 3 year intervals compared to three year cytology but no change in the number with 5 year HPV screening. The published English study predicted a fall in the number of colposcopies with a six-year HPV screening strategy but an increase with six year screening with primary partial HPV genotyping.

- The loss of utility from attending screening, having abnormal results or having a colposcopy is not well understood which generates uncertainty in overall findings for HPV compared to cytology screening strategies. This was acknowledged as a key weakness in the report describing the unpublished UK model.

Given the uncertainties inherent in modelling HPV testing to cytology based screening, there are even greater uncertainties around recommendations from studies on HPV screening intervals. Putting aside concerns about the potential limited generalisability of findings from other countries, the published studies suggest that increasing the HPV screening intervals reduces the costs of screening but reduces the potential benefits. The one English study[5] also reported this to be the case but only looked at 5 year, 6-year and 6 year followed by 10 year intervals at age 49. This study essentially found the same as the unpublished model – that shorter screening intervals had higher costs but lower gains in life expectancy. However, the differences per woman in both cost and especially QALYs and life years were very small in both the published English study and in the unpublished model with the difference in QALYs across all strategies never more than one quality adjusted life *day* over a lifetime. In addition, depending on the strategy there are differences in, for example, colposcopy rates or CIN2 detection that may be taken into account by decision makers.

It is very difficult to say with certainty which strategies are the most cost effective with such small incremental differences even without the significant uncertainties in the model results already identified. The authors of the English study for example concluded that a five or six year strategy "could be" the most cost effective strategy, although it is not clear how they reached this conclusion it appears to be based upon essentially a cost-consequences analysis based upon potential differences in, for example, colposcopy rates. If the evidence from the model supported "five or six" year screening it also supported six year screening to age 49 followed by ten year screening to age 64 depending on the weight put on QALYs over life years and other outcomes. It also supported HPV partial genotyping as the primary screening method.

Consideration of screening intervals will depend on the relative weight placed on the burden of screening, the likelihood of attendance of all women or women in different risk groups, the potential reduction in cancers and deaths and the change in the number of abnormal readings or colposcopies.

In summary, if the assumptions that have been made on HPV screening compliance in published studies represents reality, it is likely that five-year HPV screening with its longer intervals and better test parameters compared to cytology is likely to be cost saving compared to cytology screening. The Bain's model is therefore consistent with other published findings. If the disutility is not too great for abnormal screening results then it is likely that HPV screening also generates more QALYs than cytology screening. Once a decision on HPV modelling has been reached, the incremental costs and benefits of different screening intervals per woman are very small and the choice of HPV screening interval will depend on how decision makers wish to interpret model results (especially the weight they put on different utility sets and/or on non-QALY outcomes) and the assumptions they most believe on future HPV screening compliance.

Key finding two: Evidence on management of women who are HPV+/cyt- is unclear and limited by the modelling approaches chosen

Whilst four studies had made recommendations on the use of HPV genotyping and or/recall intervals in terms of the management of women who were HPV+/cyt-, the recommendations were all different. The one English study made a strong recommendation on 12 month rather than 24 month recall intervals for HPV+/cyt- women with no mention of genotyping as part of the strategy whilst studies from New Zealand and Australia – using essentially the same model – recommended some form of HPV genotype testing strategy. As stated, it was not clear in the English study how a conclusion of 12 month intervals was reached given 24 month intervals appeared to potentially be more cost effective.

Examination of the four studies revealed that, as was the case for differing primary HPV screening intervals, the results were very sensitive to the utility values chosen with the difference in QALYs regardless of the utility value set chosen in the region of quality adjusted life *hours* over a woman's lifetime. With differences that are so marginal coupled with the inherent uncertainty around compliance with HPV screening again decision makers could interpret model outputs to justify almost any strategy.

Recommendations for future modelling

It is our opinion that outside of modelling different lengths of screening, the modelling approaches that have been undertaken in the identified studies are perhaps inappropriate to answer questions about different strategies for women who test HPV+/cyt-, exit strategies for women at age 64 (or indeed any age) who are HPV+ at last screening or indeed any strategy for HPV+ women. By attempting to model both screening over a lifetime and strategies for the small percentage of women who have abnormal results the potential differential cost effectiveness of strategies for abnormal results has the potential to be drowned out by the noise and assumptions from the overall model. The reasons for this are twofold:

- Any potential cost and outcome (including QALY) differences with different strategies for dealing with abnormal results will be averaged out across all women. As the vast majority of women never have an abnormal result this heavily dilutes the cost and outcome differences between strategies for the average woman.
- The small absolute differences for the average woman from different strategies will be reduced even further by discounting. With a lifetime horizon and five year screening cycles, abnormal screenings that occur in anything other than the first screen will be discounted. With a 3.5%pa discount rate any costs and benefits of different strategies will reduce by 16% at the second screen, 30% at the third screen and 41% by the fourth screen.

To assess the cost effectiveness of different strategies for abnormal readings the correct approach – in our opinion - is to model a cohort that has an abnormal reading *only* with the model starting *at the time of the abnormal reading.* This is equally true for strategies for HPV+/cyt- at any point in the screening timeline or for women who are HPV+ at any point including the last screening.

For all future models of cervical cancer screening, utility values for colposcopy, false negatives, abnormal screenings and screening itself should be identified and collected if necessary. Close monitoring of the five year HPV screening strategy should also continue so screening compliance rates can be monitored with the model adapted with real world data rather than assumption when it becomes available.

References

1 Accetta, G.; Biggeri, A.; Carreras, G.; Lippi, G.; Carozzi, F. M.; Confortini, M.; Zappa, M.; Paci, E., Is human papillomavirus screening preferable to current policies in vaccinated and unvaccinated women? A cost-effectiveness analysis, Journal of Medical Screening 2010

2 Berkhof, J.; Coupe, V. M.; Bogaards, J. A.; van Kemenade, F. J.; Helmerhorst, T. J.; Snijders, P. J.; Meijer, C. J., The health and economic effects of HPV DNA screening in The Netherlands, International Journal of Cancer 2010

3 Burger, E. A.; Pedersen, K.; Sy, S.; Kristiansen, I. S.; Kim, J. J., Choosing wisely: a model-based analysis evaluating the trade-offs in cancer benefit and diagnostic referrals among alternative HPV testing strategies in Norway, British Journal of Cancer 2017

4 de Kok, I. M.; van Rosmalen, J.; Dillner, J.; Arbyn, M.; Sasieni, P.; Iftner, T.; van Ballegooijen, M., Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model, BMJ 2012

5 Kitchener, H. C.; Canfell, K.; Gilham, C.; Sargent, A.; Roberts, C.; Desai, M.; Peto, J., The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England: Extended follow-up of the ARTISTIC randomised trial cohort through three screening rounds, Health Technology Assessment 2014

6 Kulasingam, S. L.; Rajan, R.; St Pierre, Y.; Atwood, C. V.; Myers, E. R.; Franco, E. L., Human papillomavirus testing with Pap triage for cervical cancer prevention in Canada: a cost-effectiveness analysis, BMC Medicine 2009

7 Lew, J. B.; Simms, K.; Smith, M.; Lewis, H.; Neal, H.; Canfell, K., Effectiveness Modelling and Economic Evaluation of Primary HPV Screening for Cervical Cancer Prevention in New Zealand, PLoS ONE [Electronic Resource] 2016

8 Lew, J. B.; Simms, K. T.; Smith, M. A.; Hall, M.; Kang, Y. J.; Xu, X. M.; Caruana, M.; Velentzis, L. S.; Bessell, T.; Saville, M.; Hammond, I.; Canfell, K., Primary HPV testing versus cytology-based cervical screening in women in Australia vaccinated for HPV and unvaccinated: effectiveness and economic assessment for the National Cervical Screening Program, The Lancet Public Health 2017

9 Petry, K. U.; Barth, C.; Wasem, J.; Neumann, A., A model to evaluate the costs and clinical effectiveness of human papilloma virus screening compared with annual papanicolaou cytology in Germany, European Journal of Obstetrics Gynaecology and Reproductive Biology 2017

Appendix A Search strategy

Literature search for HPV modelling studies – November 2017

All literature searches carried out on 20th November 2017

Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

1. Papillomavirus Infections/ (22984)
2. (human adj (papilloma virus or papillomavirus)).tw. (35425)
3. HPV.tw. (36968)
4. 1 or 2 or 3 (48399)
5. ((primary or first or initial or main) adj (screen$3 or test or tests or testing or detect$3 or assessment)).tw. (35835)
6. 4 and 5 (593)
7. ((primary HPV or human papillomavirus) adj2 (screen$ or test or tests or testing)).tw. (1709)
8. HPV-DNA test$3.tw. (1136)
9. 6 or 7 or 8 (2922)
10. Uterine Cervical Neoplasms/ (73355)
11. cervical cancer.tw. (40563)
12. (cervical intraepithelial neoplasia or CIN).tw. (12533)
13. cancer of the cervix.tw. (3398)
14. 10 or 11 or 12 or 13 (90217)
15. Models, Theoretical/ (145978)
16. Models, Economic/ (9253)
17. Logistic models/ (129507)
18. Computer Simulation/ (183239)
19. Cost-Benefit Analysis/ (77128)
20. Markov Chains/ (13461)
21. Health Care Costs/ (36712)
22. Technology Assessment, Biomedical/ (9769)
23. ((Markov or mathematical or theoretical or microsimulation or simulation or economic or cost$ or clinical or benefit or effective$ or decision) adj (model$ or analy$ or evaluation$ or assessment$ or comparison$)).tw. (235286)
24. (model adj (analy$ or simulation or input$)).tw. (12555)
25. decision analy$ model$.tw. (2908)
26. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 (747071)
27. 9 and 14 and 26 (215)
28. limit 27 to yr="2005 -Current" (173)

Embase 1996 to 2017 Week 47

1. papillomavirus infection/ (10584)
2. (human adj (papilloma virus or papillomavirus)).tw. (34937)
3. HPV.tw. (41450)
4. 1 or 2 or 3 (50794)
5. ((primary or first or initial or main) adj (screen$3 or test or tests or testing or detect$3 or assessment)).tw. (37905)
6. 4 and 5 (847)
7. ((primary HPV or human papillomavirus) adj2 (screen$ or test or tests or testing)).tw. (1931)
8. HPV-DNA test$3.tw. (1477)
9. 6 or 7 or 8 (3613)
10. uterine cervix cancer/ (47257)
11. cervical cancer.tw. (44676)
12. (cervical intraepithelial neoplasia or CIN).tw. (13701)
13. cancer of the cervix.tw. (1652)
14. 10 or 11 or 12 or 13 (71484)
15. theoretical model/ (65783)
16. economic model/ (683)
17. statistical model/ (142222)
18. computer simulation/ (97496)
19. "health care cost"/ (151225)

20. biomedical technology assessment/ (10234)
21. "cost effectiveness analysis"/ (122121)
22. ((Markov or mathematical or theoretical or microsimulation or simulation or economic or cost$ or clinical or benefit or effective$ or decision) adj (model$ or analy$ or evaluation$ or assessment$ or comparison$)).tw. (233245)
23. (model adj (analy$ or simulation or input$)).tw. (14229)
24. decision analy$ model$.tw. (3914)
25. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 (721982)
26. 9 and 14 and 25 (348)
27. limit 26 to yr="2005 -Current" (287)

Cochrane Library

#1 MeSH descriptor: [Papillomavirus Infections] this term only (755)
#2 ("human papilloma virus" or "human papillomavirus"):ti,ab,kw (1439)
#3 HPV:ti,ab,kw (1519)
#4 #1 or #2 or #3 (1948)
#5 ((primary or first or initial or main) and (screen* or test* or detect* or assessment)):ti,ab,kw (149875)
#6 #4 and #5 (464)
#7 (("primary HPV" or "human papillomavirus" or "human papilloma virus") and (screen* or test*)):ti,ab,kw (711)
#8 ("HPV-DNA test*" or "HPV DNA test*"):ti,ab,kw (94)
#9 #6 or #7 or #8 (896)
#10 MeSH descriptor: [Uterine Cervical Neoplasms] this term only (2028)
#11 "cervical cancer":ti,ab,kw (1907)
#12 ("cervical intraepithelial neoplasia" or CIN):ti,ab,kw (1259)
#13 "cancer of the cervix":ti,ab,kw (46)
#14 #10 or #11 or #12 or #13 (3598)
#15 MeSH descriptor: [Models, Theoretical] this term only (983)
#16 MeSH descriptor: [Models, Economic] this term only (1565)
#17 MeSH descriptor: [Logistic Models] this term only (4920)
#18 MeSH descriptor: [Computer Simulation] this term only (1846)
#19 MeSH descriptor: [Cost-Benefit Analysis] this term only (18506)
#20 MeSH descriptor: [Markov Chains] this term only (2177)
#21 MeSH descriptor: [Health Care Costs] this term only (4693)
#22 MeSH descriptor: [Technology Assessment, Biomedical] this term only (632)
#23 ((Markov or mathematical or theoretical or microsimulation or simulation or economic or cost* or clinical or benefit or effective* or decision) and (model* or analy* or evaluation* or assessment* or comparison*)):ti,ab,kw (413383)
#24 (model and (analy* or simulation or input*)):ti,ab,kw (39283)
#25 "decision analy* model*":ti,ab,kw (340)
#26 #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 (423980)
#27 #9 and #14 and #26 Publication Year from 2005 to 2017 (273)

Search results	
Medline	173
Embase	287
Cochrane Library	273

Total	**733**

Appendix B Full data extraction tables

Paper	Country	Objective	Time Horizon	Discount rate	Model type	Model structure
Accetta 2010	Italy	To assess the cost effectiveness of HPV screening in Italy	Lifetime	3% (costs and QALYs)	Markov model with micro simulation	Markov with nine health states and death: Healthy, HPV infection (lrHPV, hr HPV 16/18, hrHPV non-16/18) pre cancer lesions (low and high severity) and cancer (local, regional, distant). Women progress though model one at a time with annual cycles with state dependent probabilities
Berkhof 2010	Netherlands	To assess the cost effectiveness of HPV screening strategies in the Netherlands	To age 100	1.5% for QALYs and 4.0% for costs.	Markov model with micro simulation	Markov with six monthly cycles. Health states not well described but can be inferred to include HPV free and HPV states with HPV low risk and high risk states. Patients can develop CIN2+ only if in the HPV high risk state but can develop CIN1 in all HPV states. Progression to CIN3 was age dependent. Patients can progress from CIN states to cancer or back to a well state.
Burger 2017	Norway	To assess the resource use (notably colposcopy) and outcomes of different HPV testing algorithms (notably the time from switching from cytology to primary HPV screening)	Lifetime	NR	Microsimulation model	Model starting at age 8 where girls/women have a probability over time of type-specific HPV incidence and clearance. This can progress to lesions and cancer which are a function of age, lesion and duration of infection. The model is stratified by HPV genotype, CIN grades and cancer stage.
de Kok 2012	Netherlands	To assess whether and it what form HPV testing is preferable to cytology in the Netherlands, including the frequency of tests	Lifetime	3% (costs and QALYs)	Microsimulation model (MISCAN)	Patient simulation model where women have a probability over time of type-specific HPV incidence and clearance. This can progress to lesions (which can clear) and cancer which are a function of age, lesion and duration of infection. The model is stratified by high risk HPV (but not specific genotype), CIN grades and cancer stage.
Kulasingham 2009	Canada	To determine the cost-effectiveness of HPV testing in three Canadian provinces	Lifetime	3% (costs and QALYs)	Markov model	Poorly described but a cohort model with yearly cycles where women can move from healthy to precancer, cancer and death. Women can move from diseased back to healthy states
Petry 2017	Germany	To evaluate the cost effectiveness of HPV screening scenarios compared to cytology	10 years (5 year screening) 6 years (3 year screening)	3% (costs only - QALYs not included in analysis)	Markov model	Model is not described beyond being a cohort model with a decision tree component. No detail of health states was provided.
Lew 2017	Australia	To evaluate different screening options - including HPV with partial genotyping - in an Australian context	Lifetime	5% (costs and benefits)	Dynamic model of HPV transmission and vaccination with Markov model for natural history of CIN and cancer	Women cycle through following states: susceptible, HPV infected (and genotype), CIN1-3 and cancer (6 stages). Women can also be vaccinated or become immune. At each screening point the screening and

Paper	Country	Objective	Time Horizon	Discount rate	Model type	Model structure
					survival coupled with a deterministic screening and treatment model	treatment model is applied with a probability of screening attendance. Modules within the screening and treatment model include a colposcopy, biopsy and treatment modules and post treatment natural history module. Model outcomes were calibrated to observed data on cancer, cancer death rates and histology and abnormality rates
Kitchener 2014	England	To determine the cost effectiveness of HPV screening compared to cytology in England	Lifetime	3.5% (Costs and QALYs)	Same model as Lew 2017	Same model as Lew 2017
Lew 2016	New Zealand	To determine the cost effectiveness of HPV screening compared to cytology in New Zealand	Lifetime	3.5% (Costs and QALYs)	Same model as Lew 2017	Same model as Lew 2017

Paper	Screening strategies considered	Screening schedules (years)	Screening start age	Screening end age
Accetta 2010	No screening, Cytology, primary HPV, Cytology followed by HPV triage, HPV followed by cytology triage	3 and 5	25	65
Berkhof 2010	Cytology, HPV with cytology triage, Combination cytology and HPV, Cytology with HPV triage.	5, 6, 7.5 and 10	30	65
Burger 2017	**1. HPV followed by cytology triage** HPV+Cyt- women are screened after 12 months. HPV+ women at this point have colposcopy. HPV- women return to previous screening period. Wait time between rescreening between 6 and 18 months was explored as was 1,2 or 3 HPV+cyt- results required before referral for colposcopy. **2. Cytology with HPV triage (current strategy)** **Cyt+(minor lesions) and** HPV+ (for high risk HPV) women are screened with cytology and HPV test after 6-12 months. High grade lesions on cytology are referred for immediate colposcopy.	3, 4, 5, 6, 8 and 10	25 (28, 31 and 34 in study but these ages are a combination of initial cytology followed by HPV)	69
de Kok 2012	Nine strategies were considered with 171 policy combinations of start age and screening frequency. The only strategies considered with results reported were cytology, HPV with cytology triage followed by a second cytology triage at 6 months for HPV+/cyt- women with HPV sensitivity of 90% and 95%	3 to 10. Only results for 5 years presented	25, 27, 30 and 32	Maximum of 70
Kulasingham 2009	Eight strategies considered for a start age of 25 (a further 19 strategies for people under 25).HPV testing only, cotesting, cytology with HPV triage, HPV with cytology triage.	1, 2, 3 and 5	25	Not clear but appears to be 70
Petry 2017	Multiple HPV strategies considered in terms of type of HPV test and cytology and also cotesting. Only results for cytology (annual) and HPV with cytology triage (3 and 5 years) are extracted	1 (cytology), 3 and 5 (HPV with cytology triage)	30	65
Lew 2017	Seven strategies considered. Only extracted were cytology (with 5 years screening), HPV with cytology triage or HPV genotyping. All HPV+ women at discharge from screening are offered colposcopy. Current cytology practice is 3 years cytology screening (5 years at 45) Various other elements of screening strategies were considered including: different options for HPV+/low grade cytology (direct colposcopy or reflex HPV); whether women are invited to attend a first screening or not; different levels of compliance with call and recall programmes	2 (cytology), 5 and 6 (HPV)	25	64-74
Kitchener 2014	HPV with cytology triage and cytology alone are two main strategies. Within the HPV triage there are sub strategies depending on the treatment pathway should a woman screen as HPV+/cyt negative. **Strategy 1**: HPV+/cyt- women are recalled for HPV with cytology triage in 24 months. HPV+/cyt- women are again recalled at 24 months	5 and 6 years and 6 years 25-49 followed by 10 years 50-64	25	64

Paper	Screening strategies considered	Screening schedules (years)	Screening start age	Screening end age
	Strategy 2: HPV+/cyt- women are recalled for HPV genotype with cytology triage in 24 months. OHR+ are referred for cytology with cyt- women again recalled at 24 months. 16/18 positive women are referred to colposcopy **Strategy 3**: Initial screen is for HPV genotype. 16/18 positive women are referred to colposcopy. HPV+/cyt- women are recalled for HPV genotype with cytology triage in 24 months. OHR+ are referred for cytology with cyt- women again recalled at 24 months and move onto a 24 month retest cycle whilst they remain OHR+/cyt-. 16/18 positive women at retest are referred to colposcopy In all strategies, recall at 12 rather than 24 months was considered.			
Lew 2016	"Cytology, HPV with cytology triage, HPV with genotyping. Within the HPV triage there are sub strategies depending on the treatment pathway should a woman screen as HPV+/cyt negative. Strategies for HPV+ women are as follows **HPV with cytology triage**: HPV+/cyt- women are recalled for HPV and cytology cotest in 12 months. HPV+ or cyt+ women are sent for colposcopy. HPV-/cyt- back onto normal screening cycle **HPV with genotyping**: HPV16/18 sent for colposcopy. OHR+ are referred for cytology with cyt- women recalled at 12 months. HPV+ women at recall are referred to colposcopy and HPV- women back onto normal screening cycle	3 (cytology), 5 (HPV)	20 (cytology), 25 (HPV)	84

Paper	Source of natural history data	Source of screening effectiveness data	Source of costs	Currency and cost year	Cost of cytology and HPV tests in model	Utilities used and source
Accetta 2010	Previous models and published literature. Model was calibrated against published literature	Published studies (HPV Ranco 2008), Cytology (Goldhaber Fiebert 2008 and Kim 2007)	Activity based costing undertaken by the authors	Euro 2006	NR	Age related utilities and cancer related quality of life from Goldhaber-Fiebert. No disutility for screening or abnormal tests
Berkhof 2010	Published literature , notably the POBASCAM study against which the model results were calibrated	Published studies (unclear which studies actually drove test sensitivity/specificity parameters)	Published studies	Euro 2007	€30.87 (HPV test). €27.51 (cytology)	Utilities in model for positive screening, CIN treatment and cancer from Mandelblatt 2002, Goldie 2004 and Maissi 2004. Positive screening utility is 0.97 which seems high and the source (Madellatt) from which this value is taken comes from a paper by Gold (1998) which is a study of utilities for a range of health conditions - none of which relate to cervical cancer or positive test results and the value chosen (0.97) which is used in the Berkhof model is not present.
Burger 2017	Published data with calibration against epidemiological data from Norway	Published literature (Ronco 2014, Arbyn 2008, Nanda (2000)	No costs reported	NR	NR	NR
de Kok 2012	Not well described. MISCAN model appears to be populated with real data on the Dutch population. Model considers other hypothetical populations with different risks of cancer, HPV and whether previous	Assumption and previously published model (Berkhof 2005)	Published Dutch cost studies	Euro 2008	€21-€33 (HPV). €26-€52 (cytology)	NR

Paper	Source of natural history data	Source of screening effectiveness data	Source of costs	Currency and cost year	Cost of cytology and HPV tests in model	Utilities used and source
	screening had occurred. The Dutch population was considered low risk					
Kulasingham 2009	Epidemiological studies in Canada	Published literature (Mayrand 2007, Mayrand 2006)	Fee schedules and published literature	CAN$ 2006		NR
Petry 2017	Published literature	ATHENA trial	Clinicians and published literature	Euro 2016	NR	NR
Lew 2017	Published literature	Published studies (unclear which studies actually drove test sensitivity/specificity parameters)	Medicare Benefits Schedule, National Hospital Cost Data Collection, Pharmaceutical Benefits Schedule	AUD 2013	Actual cost of HPV test not reported. Cost of cytology $19	Two utility sets are used to derive QALYs but neither are defined in paper or Appendix. Cross referencing of other studies (Kitchener) suggests that one set of weights (Set A) included disutilities from screening, triage, testing and management whilst the other (Set B) did not have a disutility associated with screening.
Kitchener 2014	Published literature updated with the ARTISTIC trial data	Meta-analysis (Cuxick 2006) and ARTISTIC trial results	Manufacturers of HPV tests, MAVARIC study for cytology, published studies on cancer management (Martin-Hirsch 2007 and Sherlaw-Johnson 2004)	UK£ 2010	£9.38 (HPV). £5.45 (-ve), £15.40 (Low grade), £15.56 (High grade)	Two utility sets are used to derive QALYs 1. Disutility for screening, even if negative (Simonella 2014 based on SG of 43 women from general population). 2. No disutility for screening but high disutility if abnormal screening (Insinga 2007 based on TTO with 150 women from general population) .
Lew 2016	As in Lew 2017 but calibrated to Australian/UK HPV rates and NZ age specific rates	Meta-analysis for both cytology (Arbyn 2008) and HPV (Arbyn 2012). NZ registry data was	National Cervical Screening Programme	NZD 2017/18 (published in 2016?)	$35 for HPV test. $31.10 for cytology.	Broadly as Lew 2017 but with three scenarios. 1. Disutility for screening, even if negative

Paper	Source of natural history data	Source of screening effectiveness data	Source of costs	Currency and cost year	Cost of cytology and HPV tests in model	Utilities used and source
	of lesions, cancer and cancer mortality	used for screening attendance rates				(Simonella 2014 based on SG of 43 women from general population). 2. No disutility for screening but high disutility if abnormal screening (Insinga 2007 based on TTO with 150 women from general population) . 3. As 2 but with small disutility if abnormal screening (Drolet EQ5D of 490 women with abnormal screen and 460 women with normal screen).

Paper	Incidence of CIN2+	Incidence of CIN3+	Lifetime cervical cancer risk	Mortality	Treatment of lesions prevented	Cancer prevented
Accetta 2010	NR	NR	**Five year screen** Cytology: 0.77% HPV test only: 0.64% Cytology followed by HPV triage: 0.79% HPV followed by cytology triage: 0.62% **Three year screen** Cytology: 0.65% HPV test only: 0.61% Cytology followed by HPV triage: 0.70% HPV followed by cytology triage: 0.61%	NR	NR	NR
Berkhof 2010	**Reduction compared to 5 year cytology** HPV with cytology triage: 31% Combination: 34% Cytology with HPV triage: 1% Once the screening interval reached 10 years, cancer cases with HPV with cytology triage exceeded those with 5 year cytology	NR	**Reduction compared to 5 year cytology** HPV with cytology triage: 23% Combination: 26% Cytology with HPV triage: 3% Once the screening interval reached 10 years, cancer cases with HPV with cytology triage exceeded those with 5 year cytology	NR	NR	NR
Burger 2017	NR	NR	**Reduction with HPV with cytology triage (5 year cytology with HPV triage 88.7%)** Three year screen (wait time for rescreen of HPV+cyt- women all based on 2 recalls before colposcopy. Difference with 1 or 3 recalls changed values 0.1%) 6 month: 96.8%	NR	NR	**Cancers prevented per 1,000 women with HPV with cytology triage (Cancer incidence with 5 year cytology with HPV triage 3.75)** Three year screen 6 month: 2.55 12 month: 2.68 18 month: 2.59 Five year screen

Paper	Incidence of CIN2+	Incidence of CIN3+	Lifetime cervical cancer risk	Mortality	Treatment of lesions prevented	Cancer prevented
			12 month: 96.6% 18 month: 96.4% Five year screen 6 month: 96.6% 12 month: 96.3% 18 month: 96.1% Ten year screen 6 month: 95.5% 12 month: 95.1% 18 month: 94.7%			6 month: 2.69 12 month: 2.58 18 month: 2.49 Ten year screen 6 month: 2.14 12 month: 2.02 18 month: 1.88
de Kok 2012	All results 5 year screening starting age 30 and ending age 60 and % of first primary smears with CIN2+ lesions Cytology: 0.3% HPV 90% sensitivity and two times cytology triage: 0.3% HPV 95% sensitivity and two times cytology triage: 0.3%	NR	All results 5 year screening starting age 30 and ending age 60 and cancer cases per 100,000 life years Cytology: 5.7 HPV 90% sensitivity and two times cytology triage: 5.3 HPV 95% sensitivity and two times cytology triage: 5.2	All results 5 year screening starting age 30 and ending age 60 and deaths from cervical cancer per 100,000 life years Cytology: 2.6 HPV 90% sensitivity and two times cytology triage: 2.5 HPV 95% sensitivity and two times cytology triage: 2.4	NR	NR
Kulasingham 2009	NR	NR	NR	NR	NR	All compared to no intervention and per 100,000 women HPV with cytology triage (5 years screening): 1,409 HPV with cytology triage (3 years screening): 1,978 HPV only (5 years screening): 1,559 HPV only (3 years screening): 1,784 Cytology with HPV triage (1 year screening): 1,781

Paper	Incidence of CIN2+	Incidence of CIN3+	Lifetime cervical cancer risk	Mortality	Treatment of lesions prevented	Cancer prevented
						Cotesting (5 years screening): 1,697 Cotesting (3 years screening): 1,810 Cotesting (2 years screening): 1,916
Petry 2017	**Five year screening (10 years)** Cytology alone: 0.74% HPV with cytology triage (5 year): 0.37% **Three year screening (six years)** Cytology alone: 0.51% HPV with cytology triage: 0.33%	**Five year screening (10 years)** Cytology alone: 1.41% HPV with cytology triage (5 year): 0.76% **Three year screening (six years)** Cytology alone: 1.12% HPV with cytology triage: 0.74%	NR	NR	NR	**Five year screening** HPV with cytology triage: 17,413 over ten years in population of 16 million 30-65 women compared to cytology. Reduction of 50.0% compared to cytology **Three year screening** HPV with cytology triage: 9,584 over ten years in population of 16 million 30-65 women compared to cytology. Reduction of 38.3%
Lew 2017	Change in overall cases compared to current cytology screening practice provided as range within all testing scenarios considered. Based upon total population of Australian women in 2015 **Unvaccinated** Cytology: (-2,851, -1,632) HPV with cytology triage: (-2,908, -862) HPV with partial genotyping: (-2,332, -	Change in overall cases compared to current cytology screening practice provided as range within all testing scenarios considered. Based upon total population of Australian women in 2015 **Unvaccinated** Cytology: (-677, -279) HPV with cytology triage: (492, 997) HPV with partial genotyping: (606, 935)	Range in % change in ASR for cancer compared to current cytology screening programme across all screening methodologies) Unvaccinated Cytology: (4%, 19%) HPV with cytology triage: (-20%,-5%) HPV with partial genotyping: (-21%, -12%) Vaccinated	Range in % change in ASR for cancer mortality compared to current cytology screening programme across all screening methodologies) Unvaccinated Cytology: (4%, 19%) HPV with cytology triage: (-20%,-5%) HPV with partial genotyping: (-21%, -12%) Vaccinated	NR	NR

Paper	Incidence of CIN2+	Incidence of CIN3+	Lifetime cervical cancer risk	Mortality	Treatment of lesions prevented	Cancer prevented
	564) **Vaccinated** Cytology: (-1,833, -1,049) HPV with cytology triage: (-1,941, -603) HPV with partial genotyping: (-1,892, -597)	**Vaccinated** Cytology: (-513, -251) HPV with cytology triage: (173, 532) HPV with partial genotyping: (175, 508)	Cytology: (4%, 17%) HPV with cytology triage: (-18%, -4%) HPV with partial genotyping: (-16%, -8%)	Cytology: (5%, 23%) HPV with cytology triage: (-19%, -5%) HPV with partial genotyping: (-19%, -11%)		
Kitchener 2014	All CIN2/3 for all women in England. Note only provided for 6 year screening strategy **No vaccination** Current practice: 41,309 Strategy 1: 39,464 Strategy 2: 39,850 Strategy 3: 40,585 **Vaccination** Current practice: 24,365 Strategy 1: 22,909 Strategy 2: 22,951 Strategy 3: 23,036	NR	For all women in England. Note only provided for 6 year screening strategy **No vaccination** Current practice: 0.74% Strategy 1: 0.76% Strategy 2: 0.73% Strategy 3: 0.69% **Vaccination** Current practice: 0.32% Strategy 1: 0.33% Strategy 2: 0.33% Strategy 3: 0.32%	Cervical cancer deaths per annum in England. Note only provided for 6 year screening strategy **No vaccination** Current practice: 761 Strategy 1: 741 Strategy 2: 706 Strategy 3: 663 **Vaccination** Current practice: 338 Strategy 1: 333 Strategy 2: 330 Strategy 3: 326	NR	Cervical cancer cases per annum in England. Note only provided for 6 year screening strategy **No vaccination** Current practice: 2,521 Strategy 1: 2,590 Strategy 2: 2,495 Strategy 3: 2,366 **Vaccination** Current practice: 1,064 Strategy 1: 1,104 Strategy 2: 1,096 Strategy 3: 1,083
Lew 2016	All CIN2/3 for 2.3m women in NZ. **No vaccination** Cytology: 4,308 HPV with cytology triage: 3,704 HPV genotyping: 3,995 **Vaccination**	NR	Age standardised rate (per 100,000 women) **No vaccination** Cytology: 9.1 HPV with cytology triage: 9.1 HPV genotyping: 7.7 **Vaccination**	Age standardised rate (per 100,000 women) of cervical cancer death **No vaccination** Cytology: 1.5 HPV with cytology triage: 1.5 HPV genotyping: 1.3 **Vaccination**	NR	Absolute number of cervical cancer cases per 100,000 women **No vaccination** Cytology: 160 HPV with cytology triage: 161 HPV genotyping: 140 **Vaccination**

Paper	Incidence of CIN2+	Incidence of CIN3+	Lifetime cervical cancer risk	Mortality	Treatment of lesions prevented	Cancer prevented
	Cytology: 2,645 HPV with cytology triage: 2,401 HPV genotyping: 2,527		Cytology: 5.2 HPV with cytology triage: 5.2 HPV genotyping: 4.6	Cytology: 0.8 HPV with cytology triage: 0.9 HPV genotyping: 0.7		Cytology: 92 HPV with cytology triage: 93 HPV genotyping: 83

Paper	Cytology tests	HPV tests	Colposcopies	Lifetime number of screens	Histology evaluations	Treatment for lesions	Treatment of cancer
Accetta 2010	NR	NR	NR	NR	NR	NR	NR
Berkhof 2010	NR	NR	NR	NR	NR	NR	NR
Burger 2017	NR	NR	Number of colposcopy referrals with HPV with cytology triage per 1,000 women (567 with 5 year cytology with HPV triage) all based on 2 recalls before colposcopy. Difference with 1 or 3 recalls changed values within +-500 Three year screen (wait time for rescreen of HPV+cyt- women) 6 month: 2,047 12 month: 1,547 18 month: 1,263 Five year screen 6 month: 1,662 12 month: 1,286 18 month: 1,067 Ten year screen 6 month: 1,205 12 month: 961 18 month: 814	Number of tests with HPV with cytology triage per 1,000 women (17,958 with 5 year cytology with HPV triage) Three year screen (wait time for rescreen of HPV+cyt- women all based on 2 recalls before colposcopy. Difference with 1 or 3 recalls changed values within +-1,000) 6 month: 27,258 12 month: 24,274 18 month: 22,545 Five year screen 6 month: 19,422 12 month: 17,195 18 month: 15,900 Ten year screen 6 month: 12,264 12 month: 10,878 18 month: 10,041	NR	NR	**Number of precancer treatments HPV with cytology triage per 1,000 women (197 with 5 year cytology with HPV triage)** Three year screen (wait time for rescreen of HPV+cyt- women) 6 month: 405 12 month: 352 18 month: 320 Five year screen 6 month: 361 12 month: 320 18 month: 293 Ten year screen 6 month: 293 12 month: 266 18 month: 246
de Kok 2012	NR	NR	NR	All results 5 year screening starting age 30 and ending age 60 and mean	All results 5 year screening starting age 30 and ending age 60 and % of primary	NR	NR

Paper	Cytology tests	HPV tests	Colposcopies	Lifetime number of screens	Histology evaluations	Treatment for lesions	Treatment of cancer
				number of primary screens per woman Cytology: 2.13 HPV 90% sensitivity and two times cytology triage: 2.13 HPV 95% sensitivity and two times cytology triage: 2.13	screens ending in cytology Cytology: 3.3% HPV 90% sensitivity and two times cytology triage: 4.8% HPV 95% sensitivity and two times cytology triage: 5.1%		
Kulasingham 2009	NR	NR	NR	NR	NR	NR	NR
Petry 2017	NR	NR	NR	NR	NR	NR	NR
Lew 2017	"Range in % change in annual tests compared to current cytology screening programme across all screening methodologies. (Current practice has 2.4 million tests as context) Unvaccinated Cytology: (-41%, -23%) HPV with cytology triage: (-85%,-82%) HPV with partial genotyping: (-87%, -85%) Vaccinated Cytology: (-42%, -24%) HPV with cytology triage: (-88%, -86%) HPV with partial genotyping: (-91%, -	Range in % change in annual tests compared to current cytology screening programme across all screening methodologies. (Current practice has 54,700 (unvaccinated), 31,100 (vaccinated) as context) Unvaccinated Cytology: (-21%, 81%) HPV with cytology triage: (2,061%,2,250%) HPV with partial genotyping: (2,066%, 2,255%) Vaccinated Cytology: (-24%, 149%) HPV with cytology triage: (3,591%, 3,916%) HPV with partial genotyping: (3,583%, 3,909%)	Range in % change in annual colposcopies compared to current cytology screening programme across all screening methodologies. (Current practice has 81,300 (unvaccinated), 57,900 (vaccinated) as context) Unvaccinated Cytology: (-22%, -12%) HPV with cytology triage: (-7%,20%) HPV with partial genotyping: (12%, 37%) Vaccinated Cytology: (-23%, -13%) HPV with cytology triage: (-16%, 13%) HPV with partial genotyping: (-16%, 13%)	Range in average lifetime screens compared to current cytology screening programme across all screening methodologies. (Current practice 15 as context) Unvaccinated Cytology: (9, 11) HPV with cytology triage: (7, 8) HPV with partial genotyping: (7, 8) Vaccinated Cytology: (9, 11) HPV with cytology triage: (7, 8) HPV with partial genotyping: (7, 8)	Range in % change in annual histology evaluation compared to current cytology screening programme across all screening methodologies. (Current practice has 40,000 (unvaccinated), 28,200 (vaccinated) as context) Unvaccinated Cytology: (-22%, -12%) HPV with cytology triage: (-4%,28%) HPV with partial genotyping: (17%, 46%) Vaccinated Cytology: (-23%, -13%) HPV with cytology triage: (-11%, 22%) HPV with partial	Range in % change in annual treatments (presumed of lesions) compared to current cytology screening programme across all screening methodologies. (Current practice has 21,485 (unvaccinated), 13,203 (vaccinated) as context) Unvaccinated Cytology: (-23%, -13%) HPV with cytology triage: (-21%,-9%) HPV with partial genotyping: (-17%, -8%) Vaccinated Cytology: (-26%, -16%) HPV with cytology triage: (-29%, -15%) HPV with partial genotyping: (-29%, -	NR

Paper	Cytology tests	HPV tests	Colposcopies	Lifetime number of screens	Histology evaluations	Treatment for lesions	Treatment of cancer
	89%) "				genotyping: (-11%, 22%)	15%)	
Kitchener 2014	Total tests in England pa. Note only provided for 6 year screening strategy **No vaccination** Current practice: 3,703,772 Strategy 1: 636,790 Strategy 2: 636,161 Strategy 3: 564,796 **Vaccination** Current practice: 3,663,477 Strategy 1: 493,864 Strategy 2: 493,749 Strategy 3: 486,707	Total tests in England pa. Note only provided for 6 year screening strategy **No vaccination** Current practice: 245,330 Strategy 1: 2,255,505 Strategy 2: 2,251,914 Strategy 3: 2,244,887 **Vaccination** Current practice: 210,687 Strategy 1: 2,272,954 Strategy 2: 2,272,615 Strategy 3: 2,271,942	Total colposcopies in England pa. Note only provided for 6 year screening strategy **No vaccination** Current practice: 128,254 Strategy 1: 110,393 Strategy 2: 123,140 Strategy 3: 154,754 **Vaccination** Current practice: 89,848 Strategy 1: 72,943 Strategy 2: 74,112 Strategy 3: 77,048	NR	NR	NR	NR
Lew 2016	NR	NR	NR	NR	NR	NR	NR

Paper	Total discounted cost of strategies	Total QALYs of strategies	ICERs
Accetta 2010	**Five year screen** Cytology: €120 HPV test only: €176 Cytology followed by HPV triage: €113 HPV followed by cytology triage: €136 **Three year screen** Cytology: €160 HPV test only: €228 Cytology followed by HPV triage: €149 HPV followed by cytology triage: €175	**Five year screen** Cytology: 29.42631 HPV test only: 29.42958 Cytology followed by HPV triage: 29.42594 HPV followed by cytology triage: 29.42991 **Three year screen** Cytology: 29.42822 HPV test only: 29.43042 Cytology followed by HPV triage: 29.42803 HPV followed by cytology triage: 29.43048	Not reported in study. Calculated with five year HPV followed by cytology test triage as reference case **Five year screen** Cytology: extendedly dominated HPV test only: dominated Cytology followed by HPV triage: extendedly dominated **Three year screen** Cytology: dominated HPV test only: €180,392 Cytology followed by HPV triage: dominated HPV followed by cytology triage: €68,421
Berkhof 2010	**Total discounted costs per woman compared to five year cytology only (includes screening, diagnoses and treatment and indirect costs)** **Five year screen** HPV followed by cytology triage: €79.7 Cytology followed by HPV triage: €0.1 Combined: €181.9 **Six year screen** HPV followed by cytology triage: €30.0 Cytology followed by HPV triage: -€34.2 Combined: €114.8 **7.5 year screen** HPV followed by cytology triage: -€17.7 Cytology followed by HPV triage: -€66.6 Combined: €53.2	Not reported although figure in study suggests that QALYs increase as screening interval shortens and that combination testing always has higher QALY gain than HPV followed by cytology triage	**ICER compared to five year cytology only (includes screening, diagnoses and treatment and indirect costs)** **Five year screen** HPV followed by cytology triage: €9,305 (most cost effective strategy) Cytology followed by HPV triage: €3,955 Combined: €16,303 **Six year screen** HPV followed by cytology triage: €6,138 Cytology followed by HPV triage: QALY loss (no ICER calculated) Combined: €12,444 **7.5 year screen** HPV followed by cytology triage: €878 Cytology followed by HPV triage: QALY loss (no ICER calculated) Combined: €11,088 **Ten year screen** HPV followed by cytology triage: QALY loss (no ICER calculated) Cytology followed by HPV triage: QALY loss (no ICER

Paper	Total discounted cost of strategies	Total QALYs of strategies	ICERs
	Ten year screen HPV followed by cytology triage: -€61.8 Cytology followed by HPV triage: -€90.6 Combined: -€7.5		calculated) Combined: €22,452
Burger 2017	NR	NR	NR
de Kok 2012	NR	NR	NR
Kulasingham 2009	NR	NR (ICERs are for life years	ICERs not reported for all strategies and only efficiency frontier shown. For the whole of Canada, the ICER (cost/LY) for HPV with cytology triage with 5 year screening was $6,720. For 3 year screening the ICER was $24,257
Petry 2017	**Five year screening** Cytology: €176.9m pa (population of 16m) HPV with cytology triage: €117.0m pa (34% reduction) **Three year screening** Cytology: €205.6m pa (population of 16m) HPV with cytology triage: €203.9m pa (1% reduction)	NR	NR
Lew 2017	Unvaccinated Current practice: $384 per person Cytology: $242-$294 per person HPV with cytology triage: $260-$310 per person HPV with partial genotyping: $274-$323 per person Vaccinated Current practice: $325 per person Cytology: $193-$243 per person HPV with cytology triage: $202-$243 per person HPV with partial genotyping: $207-$248 per person	Both sets included disutilities with having cancer. QALY gains over current practice across strategies are small with Set A (gains in region of 0.005) and for HPV strategies only negative with Set B (losses in region of 0.002).	NR

Paper	Total discounted cost of strategies	Total QALYs of strategies	ICERs
Kitchener 2014	Total lifetime costs per woman **No vaccination** Current practice: £159 <u>Strategy 1</u> Six year screening (base case): £132 Six year (25-49) & ten year (50+) screening: £128 12 month rather than 24 month recall: £147 Five year screening: £143 <u>Strategy 2</u> Six year screening (base case): £135 Six year (25-49) & ten year (50+) screening: £131 12 month rather than 24 month recall: £152 Five year screening: £146 <u>Strategy 3</u> Six year screening (base case): £144 Six year (25-49) & ten year (50+) screening: £140 12 month rather than 24 month recall: £161 Five year screening: £156 **Vaccination** Current practice: £129 <u>Strategy 1</u> Six year screening (base case): £101 Six year (25-49) & ten year (50+) screening: £97 12 month rather than 24 month	Whilst all strategies generated a QALY gain per woman these changes were only seen in the third decimal place with the difference between the strategy with the lowest QALYs and highest QALYs being 0.0026QALYs (less than 1 quality adjusted life day over a woman's lifetime). These results only apply when utility weightings including a loss for screening were included. With utility weightings with no loss for screening but a significant loss for abnormalities all HPV strategies produced a QALY loss compared to cytology based screening	ICERs not calculated but (with the exception of those listed previous with QALY losses compared to current practice) all strategies and variants were dominant (cost and QALY saving) compared to current practice when a utility decrement for screening was applied.

Paper	Total discounted cost of strategies	Total QALYs of strategies	ICERs
	recall: £112 Five year screening: £112 <u>Strategy 2</u> Six year screening (base case): £102 Six year (25-49) & ten year (50+) screening: £97 12 month rather than 24 month recall: £113 Five year screening: £112 <u>Strategy 3</u> Six year screening (base case): £102 Six year (25-49) & ten year (50+) screening: £98 12 month rather than 24 month recall: £113 Five year screening: £113 Strategy 2: 74,112 Strategy 3: 77,048		
Lew 2016	**No vaccination** Cytology: $31.7m HPV with cytology triage: $28.7m HPV genotyping: $30.4m **Vaccination** Cytology: $25.9m HPV with cytology triage: $22.5m HPV genotyping: $22.7m	In all cases QALY changes were small (less than 0.008). HPV genotyping always had the highest QALY gain for both vaccinated and unvaccinated women if only a small disutility for abnormal screening applied. If utility weights with a high disutility for abnormal screening applied then cytology always produces a QALY gain	In terms of cost/LY saved and cost/QALY HPV genotyping was the most cost effective strategy (saving LY and cost) with HPV testing with cytology triage being less costly but less effective than 3 year cytology screening. Actual ICERs not reported as HPV was either always cost saving compared to cytology

Paper	Summary of cost effectiveness results	Summary of deterministic SA	Summary of PSA	Study recommendation	Limitations
Accetta 2010	Current strategy of primary HPV every three years is dominated by primary HPV with cytology triage every three years. Five year screening would be less expensive but with slightly worse outcomes.	Study explored screening versus vaccination and the sensitivity analysis only explored vaccination efficacy	NR	Three year screening with HPV with cytology triage	Outside of vaccine efficacy, the authors acknowledged the model is based upon Italian costs and more extensive parameter searching may have resulted in different parameter values
Berkhof 2010	Strategies with a screening interval over 7.5 years were not cost effective with a willingness to pay threshold of €20,000/QALY. The optimal strategy was five year screening with HPV followed by cytology triage	The finding that the optimal strategy was five year screening with HPV followed by cytology triage results were insensitive to changes in treatment and test costs, discount rates and HPV test sensitivity and specificity considered by the authors	Not reported but did show results from various calibration settings for CIN2+ detection rates of the model that showed a large variation in the effect of HPV screening on cancer.	Five year screening with HPV with cytology triage	The variation in results based upon calibration settings were stated as a limitation, as was that the model did not account for natural immunity after infection.
Burger 2017	Lengthening the time between screenings did have impact on the cancer incidence rate with more frequent screening reducing the cancer rate. However, the most important factor was starting screening at age 25. Different intervals between rescreening of HPV+/cyt- women or altering the point at which colposcopy occurs for these women made little impact on the cancer incidence rate	Lengthening the time between screenings did have impact on the cancer incidence rate with more frequent screening reducing the cancer rate. However, the most important factor was starting screening at age 25. Different intervals between rescreening of HPV+/cyt- women or altering the point at which colposcopy occurs for these women made little impact on the cancer incidence rate. As such, to minimise the increase in colposcopy that comes with HPV testing more retests for HPV+/cyt- women and/or longer intervals between retests should be considered.	NR	HPV-based screening among unvaccinated women should start at age 25 with an appropriate use of cytology triage to control colposcopy referrals. No recommendation was made on the frequency of testing	No trade off of resource use and benefit was considered. No account of anxiety in longer waiting times between retesting HPV+/cyt- women was taken. The findings on colposcopy were dependent on the initial strategy. HPV genotyping was not considered. There is an absence of data on future screening behaviour and loss to follow up.
de Kok 2012	The cost effectiveness results of strategies were not presented. The summary of the results states in most scenarios primary HPV screening is the preferred	HPV was preferred in all scenarios except if the cost of cytology were low or HPV prevalence was high with a high HPV test cost	NR	Where screening is well controlled, European countries should switch from cytology to HPV screening	Did not consider strategies that varied by age. Personal characteristics were not varied. The model assumed that people who do not go to screening are

Paper	Summary of cost effectiveness results	Summary of deterministic SA	Summary of PSA	Study recommendation	Limitations
	scenario in women over 30				at higher risk and if this is not the case the cost effectiveness results may not hold. No loss of utility from a positive HPV test was considered.
Kulasingham 2009	HPV testing every three years followed by cytology triage may be more effective and less costly than cytology screening alone	Lower discount rates (<2.0%) would favour 5 year screening	PSA takes into account strategies with start age of 18 so difficult to interpret. However, above a WTP threshold of approximately $25,000/LY three year screening would be more likely to be cost effective than five year screening	HPV with cytology triage from age 25 with three year screening	Lack of test performance by age, vaccination was not included
Petry 2017	Screening strategies for HPV results in fewer cancers at a lower cost than cytology alone. Screening of HPV at intervals less than five years does result in more cancers detected but at a higher cost than five year screening	NR	NR	No recommendation for a specific HPV screening strategy	Model does not include costs of treatment outside of initial cancer treatment (e.g. excludes rehabilitation and physio) and does not included costs associated with recurrence. However, this limitation favours cytology and so should not prejudice results. Vaccination was also not included in the model and failure to attend screening was not considered.
Lew 2017	The authors conclude that HPV testing every five years with partial genotyping or cotesting with cytology were the most effective. Sending those with HPV16/18 for colposcopy and other genotypes for reflex cytology was described as "one of the most cost-effective" strategies. Whilst the analysis is unambiguous that all strategies will result in lower cost and HPV strategies are likely to dominate	Range of SA undertaken and scenario analysis undertaken including: Six year HPV screening: Reduces costs for all HPV strategies between 7.5% and 8.5% (unvaccinated) and 9.4% and 10.6% (vaccinated). Increases cancer incidence or cancer mortality by between 3.1% and 4.0% (vaccinated and unvaccinated)	The PSA results suggest that across the HPV strategies considered genotyping and non-genotyping show little difference in costs or benefits with life years gains from the best performing genotype strategy and worse performing non-genotype strategy being different by less than 0.0005LY (equivalent to approximately 4 hours over a lifetime) with a discounted lifetime cost	HPV testing every five years with partial genotyping and direct colposcopy if 16/18	Results are sensitive to screening assumptions such as the return rate. The authors acknowledge there is little clinical evidence on the outcomes with referral straight to colposcopy with HPV 16/18 or reflex cytology.

Paper	Summary of cost effectiveness results	Summary of deterministic SA	Summary of PSA	Study recommendation	Limitations
	non-HPV strategies (at least if only life years and not QALYs are considered) there is no full incremental analysis of strategies, QALY gains are small across strategies and may be negative for some HPV strategies and a wide range of different scenarios were undertaken making it difficult to isolate the actual effect of different aspects of strategies.		difference between the most expensive and least expensive strategies of $50 per person		
Kitchener 2014	HPV testing is a cost effective strategy compared to cytology. Whilst most of the strategies considered were cost and QALY saving, they all resulted in greater numbers of colposcopies and biopsies in unvaccinated women. The QALY gains per woman were small with any strategy although primary HPV genotype testing only appears to be an efficient strategy in vaccinated women.	Results were sensitive to test characteristics of HPV and cytology and level of compliance with 12 or 24 month follow up for those recalled with HPV+/cyt- results	PSA was only conducted around sexual behaviour assumptions. The PSA did not indicate that the assumptions matter to the overall model results	HPV testing every five to six years with no more than 12 months recall for women who are HPV+/cyt-	A number of assumptions had to be made about future population behaviour and future costs of HPV screening for which there is little evidence.
Lew 2016	At a WTP threshold of $50,000/LY, in both unvaccinated and vaccinated women HPV genotyping was the most cost effective strategy. When QALYs were considered (although detailed findings not presented in body of report) findings are reported to vary widely. If disutility for screening and/or a minor disutility for abnormal findings are considered, then HPV genotyping remains the cost	Only undertaken for HPV with genotyping. Total costs were found to be sensitive to the cost of cytology and HPV tests, test characteristics of HPV and aggressiveness of natural history of HPV. Life years were found to be sensitive to the aggressiveness of natural history of HPV and adherence to screening strategy	NR	Primary HPV with genotyping	Parameters related to future screening practice were assumptions, notably the cost of the HPV test with HPV screening not being cost saving if the cost of the HPV test increased from the base case of $35 to $40. It is not commented on in the paper that the current cost of the test is £43 and they have assumed it would be $35 if HPV screening was the strategy adopted. How this assumption is derived is not described.

Paper	Summary of cost effectiveness results	Summary of deterministic SA	Summary of PSA	Study recommendation	Limitations
	effective choice. If there is no disutility from screening itself but a major disutility from abnormal findings then all HPV strategies are less effective than cytology screening				

UK National Screening Committee

Cervical cancer screening - HPV self sampling

External review of key questions relating to screening programme implementation for the UK National Screening Committee (UK NSC)

Version: Second Draft

Version date: 14th March 2017

Costello Medical Consulting Ltd.

March 2017

The UK NSC advises Ministers and the NHS in all 4 UK countries about all aspects of screening policy. Its policies are reviewed on a 3 yearly cycle. Current policies can be found in the policy database at http://legacy.screening.nhs.uk/screening-recommendations.php and the policy review process is described in detail at https://www.gov.uk/guidance/evidence-and-recommendations-nhs-population-screening#evidence-review-process

Template v2.0, August 2016

1.1 Objectives

1.1.1 Use of self-sampling to improve screening uptake

Cervical screening coverage has fallen in recent years, especially amongst younger women. Since 2004, uptake has dipped below 80% and 5-year coverage among women aged 25 to 29 is now below two-thirds.[19] The NCSP would like to pilot the use of self-sampling as a means of improving screening uptake.

This review was part of a larger piece of work and the sections relevant to self sampling have been extracted from the larger document.

Question 4 aims to establish whether self-collected specimens are of comparable accuracy to clinician-collected specimens, and **Question 5** investigates whether inviting unscreened women to return a self-collected specimen increases overall uptake of screening.

1.1.2 Identified evidence

The specific questions addressed in this review are shown in Table 1 below, along with the relationship of each question to the UK NSC's Screening Criteria and the number of studies that were identified as providing relevant evidence for each question.

Table 1 Key questions for the evidence summary, and relationship to UK NSC screening criteria

Criterion		Key questions	# studies included
THE TEST			
4	There should be a simple, safe, precise and validated screening test.	Question 4: What is the accuracy of HPV testing in self-collected specimens?	1 systematic literature review (SLR) 1 primary study
6	The test, from sample collection to delivery of results, should be acceptable to the target population.	Question 5: Does self-collection of vaginal specimens increase uptake of cervical screening?	1 SLR 10 primary studies

1.2 Methods

The current review was conducted by Costello Medical Consulting, in collaboration with the UK National Screening Committee. Database searches were conducted on 20th October 2016.

2 Synthesis of evidence

2.1 Overall results

Database searches yielded 204 results, of which 12 records were judged to be relevant to this review. An additional record was pre-specified for inclusion at the start of the review so 13 articles were ultimately included.

A study-level summary of data extracted from each included publication is presented in Appendix 4. Results of the quality assessments are also presented in Appendix 4.

2.2 Use of self-sampling to improve screening uptake

2.2.1 Use of self-sampling to improve screening uptake

2.2.1.1 Question 4 – What is the accuracy of HPV testing in self-collected specimens?

Criterion 4 of the UK NSC Screening Criteria states that: 'There should be a simple, safe, precise and validated screening test.'

This review looked for prospective studies which directly compared the accuracy of HPV testing on clinician-collected or self-collected samples.

2.2.1.1.1 Description of the evidence

One systematic review and meta-analysis, Arbyn 2014,[35] and a single additional study[36] have been identified which assess the accuracy of HPV testing on self-collected samples.

The systematic review and meta-analysis examined 36 studies, with a combined population of 154,556 women, considering the comparative accuracy of self-collected and clinician-collected self-sampling for HPV testing. The review considered the accuracy of screening in 3 population groups: women in a 'healthy screening population' attending for cervical cancer screening, high-risk women, and women in a 'follow-up' population who had been referred for colposcopy. The population meeting the eligibility criteria for this review is the 'healthy screening population'. This group was assessed in 16 of the 36 studies included in Arbyn 2014, with some outcomes having been reported separately for this population.[35]

The additional primary study identified in this review, Stanczuk 2016, assessed the performance of a polymerase chain reaction (PCR)-based assay for the detection of HPV in self-collected vaginal and urine samples. The study investigated a cohort of 5,318 women attending routine screening in a primary care setting.[36]

2.2.1.1.2 Quality assessment

An assessment of the methodological quality of the SLR using the AMSTAR checklist demonstrated overall good quality, with only one of the 11 checklist questions not addressed.[35]

Arbyn 2014 assessed the quality of its studies using the QUADAS-2 checklist and reported overall moderate to good quality in its studies (Table 2).[35] An assessment of the quality of the Stanczuk 2016 publication,[36] using the same checklist, determined a low risk of bias and low concerns of applicability relating to participant selection, index tests and reference standards in the study. The risk of bias with regards to participant flow was assessed as being high because not all participants had available results for each test, and not all participants received colposcopy (Table 2).

Table 2 Quality assessment (QUADAS-2) of studies included in Arbyn 2014 and Stanczuk 2016

Study reference	Arbyn 2014[35]	Stanczuk 2016[36]
PARTICIPANT SELECTION		
Risk of bias	Moderate	Low
Concern about applicability		Low
INDEX TESTS		
Risk of bias	Low	Low
Concern about applicability		Low
REFERENCE STANDARD		
Risk of bias	Low	Low
Concern about applicability		Low
PARTICIPANT FLOW		
Risk of bias	Moderate	High

Participant selection

A low risk of bias with regards to participant selection is reported for 53% of all included studies in the SLR and meta-analysis, while a medium risk of bias is reported for 44%. There are concerns regarding the applicability of this SLR to this review question due to the inclusion of participants in a 'high-risk screening population' and women who had been called for follow-up. These populations do not fit the eligibility criteria for this review; however, absolute accuracy values are reported separately for each group, including for women in the healthy screening population of interest to this review. Furthermore, while some outcomes are reported as pooled results across all included studies, it was demonstrated that the variability in results across studies in different populations was very low, which reduces concerns about the risk of bias in the meta-analysis.[35]

A low risk of bias for participant selection was determined for Stanczuk 2016. The study assessed women attending routine screening in Scotland, excluding only women who had previously been diagnosed with CIN2+, which is considered to be an appropriate exclusion. The screening population in Scotland is slightly younger than in England, with 97% of participants aged between 20 and 59 years old, however, the risk of bias associated with this is judged to be very low.[36]

Index test

No studies in the SLR were assessed to have a high risk of bias with regards to the reporting or execution of index tests. The approach was considered adequate in 72% of studies and in 28% it was unclear.[35]

The screening in Stanczuk 2016 was undertaken in accordance with the UK Cervical Cancer Screening Programme and, therefore, is applicable to this review. The publication does not report the threshold for a positive HPV result, however, this is not considered a serious quality concern and generally the execution of index test was considered appropriate, resulting in an overall low risk of bias.[36]

Reference standard

The quality of test verification with a reference standard is reported in Arbyn 2014 to be good in 89% of studies, moderate in 8% of studies and possibly problematic in one study. The SLR required studies to have used either colposcopy, considered to be the gold standard, or biopsy as the reference standard and to assess either CIN2+ or CIN3+ as the target abnormality. These eligibility criteria are aligned with those applied in this review, and as a result, the studies identified are highly applicable.[35]

The overall risk of bias associated with the reference standard was assessed to be low in Stanczuk 2016. The study uses colposcopy as the reference standard and detects both CIN2+ and CIN3+. The publication does not report whether the reference standard results are interpreted without knowledge of the index tests results, which has potential to cause bias in test verification. However, in line with the approach taken in Arbyn 2014 with regards to unclear blinding, this domain was judged to have overall low risk of bias.[36]

Participant flow

A moderate risk of bias associated with participant flow was reported in Arbyn 2014 on assessment of all included studies in the SLR. The delay between self-sampling, clinician-sampling and verification was determined to be short (<6 months) in 69% of included studies, unreported in 9% and long in 6% of studies. Partial verification was avoided in 78% of studies and differential verification avoided in all but one study.[35] Arbyn 2014 noted that when the delay between tests was not reported the sensitivity was significantly lower than when it was clearly reported.[35]

The recall time for colposcopy referral was not reported in Stanczuk 2016, and additionally, not all patients enrolled in the trial received a reference standard or were included in analyses. Due to these concerns the study was determined to have a high risk of bias associated with participant flow.[36]

2.2.1.1.3 Results

Accuracy of testing on self-collected and clinician-collected samples were reported as pooled results from the Arbyn 2014 meta-analysis,[35] and additionally in Stanczuk 2016.[36] Study-level details of these results are presented in Appendix 4, Table 15. A summary of the results is presented below (Table 3).

Table 3 Summary of accuracy results from the Arbyn 2014 meta-analysis and Stanczuk 2016

	Arbyn 2014[a, 35]				Stanczuk 2016[36]		
		Clinician-collected			Self-collected		
	Self-collected	HPV	Cytology ASC-US+	Cytology LSIL+	Vaginal sample, n=5208	Urine sample, n=5003	Clinician-collected, n=5299
CIN2+							
Studies, n	16	16	12	8	NA		
Sensitivity (95% CI)	76% (69 to 82)	91% (87 to 94)	83% (75 to 89)	71% (66 to 76)	94.6% (90.7 to 98.5)	63.1% (54.6 to 71.1)	97.7% (95.0 to 100)
Specificity (95% CI)	86% (83 to 89)	88% (85 to 91)	91% (87 to 94)	97% (97 to 98)	85.4% (84.4 to 86.3)	59.8% (89.0 to 90.7)	87.3% (86.4 to 88.2)
LR+	NR	NR	NR	NR	6.48	1.57	7.69
LR-	NR	NR	NR	NR	0.054	0.62	0.026
CIN3+							
Studies, n	8	8	6	5	NA		
Sensitivity (95% CI)	84% (72 to 92)	95% (91 to 97)	91% (85 to 95)	78% (72 to 85)	95.8% (91.1 to 100)	50.7% (39.1 to 62.3)	98.6% (95.9 to 100)
Specificity (95% CI)	87% (84 to 90)	89% (87 to 92)	89% (86 to 91)	97% (96 to 97)	84.8% (83.8 to 85.8)	89.7% (88.8 to 90.5)	86.4% (85.5 to 87.3)
LR+	NR	NR	NR	NR	6.30	4.92	7.25
LR-	NR	NR	NR	NR	0.045	0.49	0.016

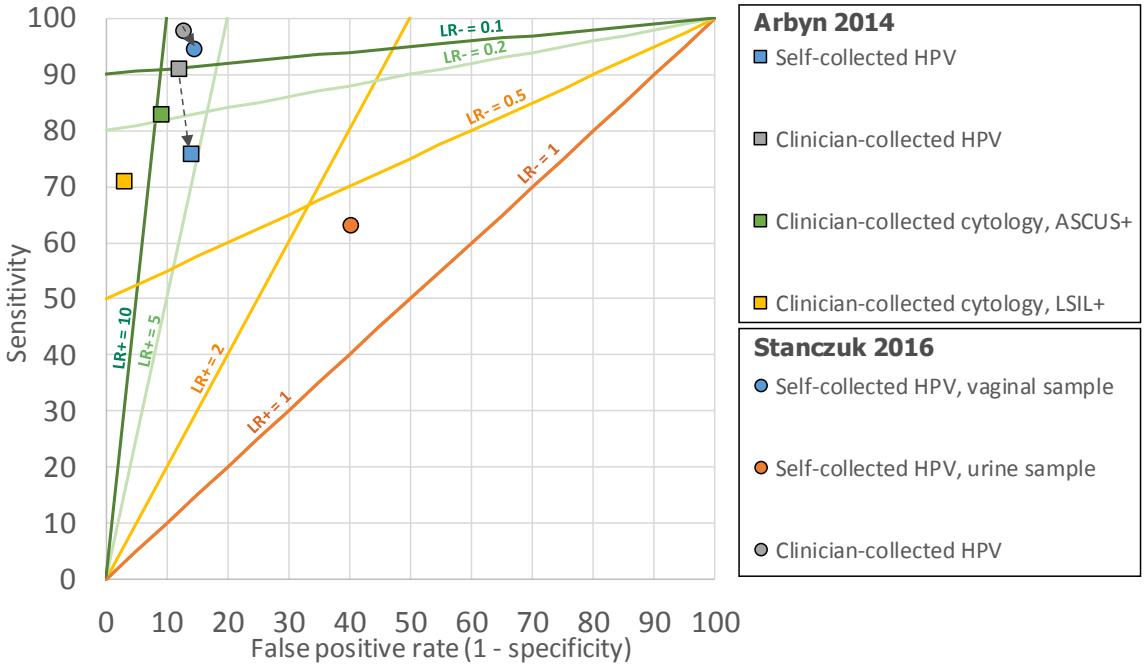
ASC-US, Atypical Cells of Undetermined Significance; CIN, Cervical Intraepithelial Neoplasia; LSIL, Low grade Squamous Intraepithelial Neoplasia

[a] Pooled values across studies in the systematic review enrolling participants considered part of a healthy screening population

Accuracy

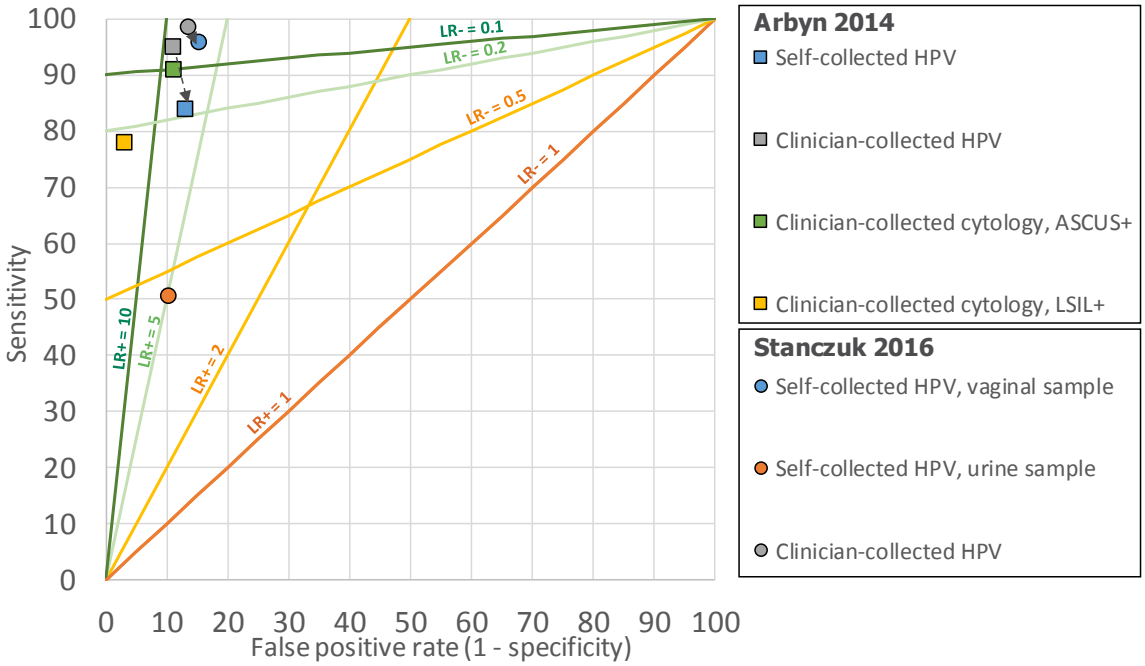
The following figures present the accuracy for self-collected and clinician-collected HPV samples pooled across the studies considering a healthy screening population in the Arbyn 2014 meta-analysis.[35] The figures additionally present the absolute accuracy of HPV testing on the 2 self-collection methods and clinician-collected samples determined in Stanczuk 2016.[36] The accuracy is presented separately for the detection of CIN2+ (Figure 1) and CIN3+ (Figure 2). The pooled clinician-collected cytology results reported in Arbyn 2014 are presented alongside the self-collected HPV test results to allow comparison to current clinical practice.[35]

Figure 1 Accuracy of screening methods for the detection of CIN2+



ASC-US, Atypical Cells of Undetermined Significance; CIN, Cervical Intraepithelial Neoplasia; HPV, Human Papilloma Virus; LSIL, Low grade Squamous Intraepithelial Neoplasia

Figure 2 Accuracy of screening methods for the detection of CIN3+



ASC-US, Atypical Cells of Undetermined Significance; CIN, Cervical Intraepithelial Neoplasia; HPV, Human Papilloma Virus; LSIL, Low grade Squamous Intraepithelial Neoplasia

Relative accuracy

The relative sensitivity and specificity results for detection of CIN2+ (Figure 3 and Figure 4) and CIN3+ (Figure 5 and Figure 6) with HPV testing are presented below. As discussed previously, the results from Arbyn 2014 were pooled across all 36 included studies, including those in high-risk and follow-up populations. A small variability in the results between different groups was demonstrated, and as a result, a low risk of bias is associated with this meta-analysis.[35] The results for Stanczuk 2016 present the accuracies of the 2 self-collection methods relative to the clinician-collected sample.[36]

Figure 3 Relative sensitivity of screening methods for detection of CIN2+

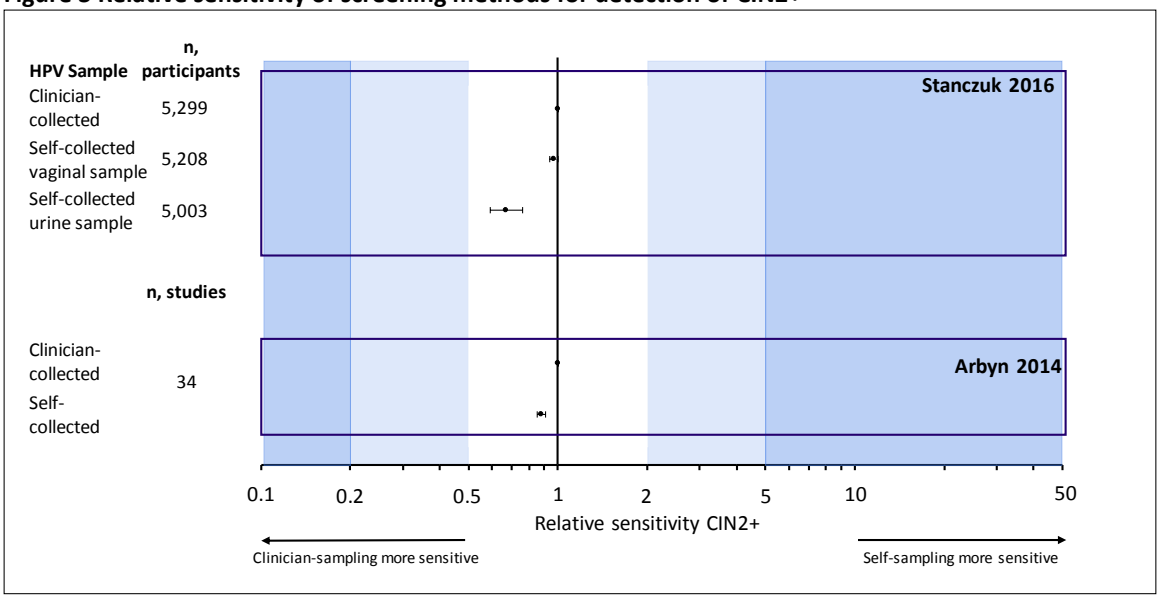


Figure 4 Relative specificity of screening methods for detection of CIN2+

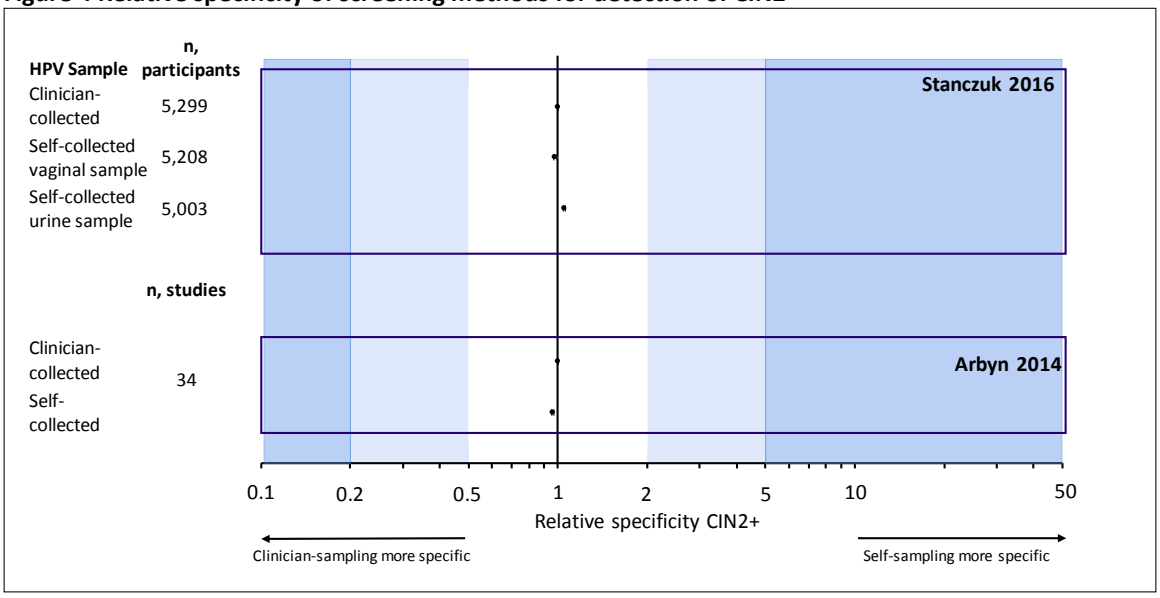


Figure 5 Relative sensitivity of screening methods for detection of CIN3+

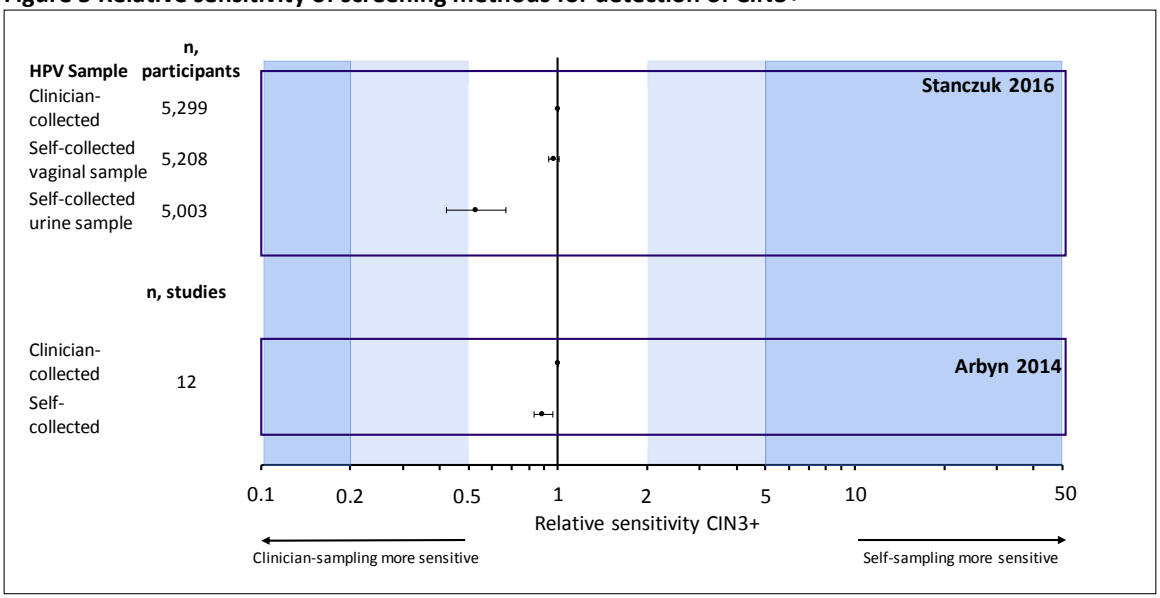
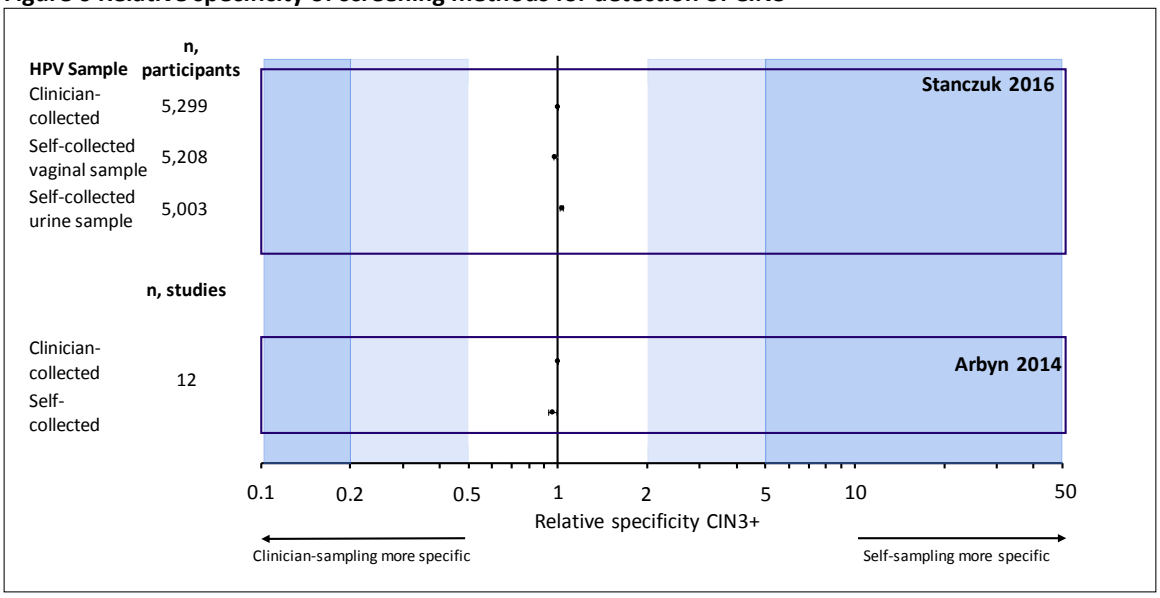


Figure 6 Relative specificity of screening methods for detection of CIN3+



The pooled relative accuracy results presented in the meta-analysis demonstrate that the sensitivity and specificity of HPV testing for detection of CIN2+ and CIN3+ on self-collected samples compared to clinician-collected samples are statistically significantly lower, but the difference is small. The pooled absolute sensitivity reported in Arbyn 2014 for the detection of CIN2+ was 76%, with results from individual studies ranging from 51% to 93%, within studies in a healthy screening population. The meta-analysis additionally reported a pooled absolute specificity of 86%, ranging from 67% to 93%. For the detection of CIN3+ the pooled results demonstrated a sensitivity of 84%, ranging from 63% to 94%, and a specificity of 87%.[35] While broad ranges in accuracy results are presented, Arbyn 2014 reported that issues relating to study design or reporting quality have not altered the accuracy of results in the meta-analysis. A single exception is that significantly lower sensitivity and specificity values are reported when

the delay between tests is not clearly reported. Furthermore, the variation between study results can partially be rationalised by heterogeneity in the HPV test used.

The absolute sensitivity values reported in Stanczuk 2016 for HPV testing on self-collected vaginal samples are within the range reported for similar studies in Arbyn 2014, with sensitivity for detection of CIN2+ and CIN3+ of 94.6% (95% CI 90.7% to 98.5%) and 95.8% (95% CI 91.1% to 100%) respectively. The absolute specificity observed was in agreement with the values from the meta-analysis. HPV testing on self-collected vaginal samples for the detection of both CIN2+ and CIN3+ was considered as sensitive and as specific as the testing on clinician-collected samples.[36] The relatively high sensitivity values may be explained by the use of a PCR detection method in Stanczuk 2016, which is known to be more analytically sensitive than other tests, such as the Hybrid Capture 2 (HC2) HPV test which was used in the majority of studies identified in Arbyn 2014.[37]

Overall, the accuracy results reported in Arbyn 2014 should be considered at low risk of bias. The SLR was well conducted, and identified studies which are highly relevant to this review in terms of study population and reference testing.[35] The results are also supported by an additional study identified in this review.[36]

Further points to note are considerations made in Arbyn 2014 regarding self-sampling devices and HPV test methods. No significant differences were observed between the different self-sampling devices used across included studies. The assay used in the vast majority of studies identified in the SLR was HC2. Using HC2 as a reference, generally no significant differences were observed to other HPV tests. There were just 2 exceptions: a significantly higher sensitivity using a 'MALDI-TOF' test; and significantly higher specificity using an APTIMA HPV test.[35]

2.2.1.1.4 Evidence summary

This review considers the accuracy of HPV testing on self-collected samples based on evidence from a good quality SLR[35] which has been demonstrated to have a good applicability to this review and results from one further primary study.[36]

Evidence across these publications suggests that the accuracy of HPV testing in self-collected samples is lower than in clinician-collected samples, but not substantially lower, and could be considered an appropriate alternative to clinician-sampling for women not attending primary screening.[35, 36] The high relative accuracy results reported in Stanczuk 2016 on self-collected vaginal samples highlight the potential for optimisation of testing by varying the HPV assay,[36] a factor which could warrant further investigation.

Analysis of evidence relevant to criterion 4 – Accuracy of HPV testing on self-collected specimens

Quantity: Overall the evidence identified comprises one SLR[35] and one additional primary study in a cohort of 5,318 women.[36] The SLR reports 36 primary studies, 16 of which are of high relevance to this review. The total number of studies represent a reasonable evidence base to consider the accuracy of HPV testing on self-collected samples.

Quality: An assessment of the methodological quality of the SLR demonstrated that it had overall good quality, with only one question in the AMSTAR checklist not having been addressed.[35] The SLR reported the quality of its included studies as moderate to good and an assessment of the quality of Stanczuk 2016 showed it to have a generally low risk of bias to this review question.[36] The only exception was that not all patients enrolled in the study received a reference standard or were included in the analyses, potentially biasing the accuracy results.

Applicability: The eligibility criteria for the inclusion of studies in Arbyn 2014 were generally well aligned to those applied in this review.[35] One concern regarding the study populations included in the SLR has been considered. Studies examining both women who are considered at high-risk, or who had been invited to follow-up cytology were included, while the population of interest in this review is a healthy screening population. However, some outcomes were presented separately for a low-risk population, and it was demonstrated that the variability between the populations was low, leading to a low risk of bias in results pooled across all included studies. Overall, the results from the SLR are considered applicable to this review. An assessment of applicability of Stanczuk 2016 using the QUADAS-2 checklist also demonstrated a low risk of bias.[36]

Consistency: A relatively broad range of absolute accuracy values for the detection of CIN2+ and CIN3+ in self-collected samples were reported in Arbyn 2014. However, to an extent, this result can be rationalised by the difference in HPV test methodology in the included studies and it was demonstrated that, with one exception, issues of study quality did not impact on accuracy.[35] The absolute accuracy values reported in Stanczuk 2016 fall within this range and support the consistency of the result presented from the meta-analysis.[36]

Conclusion

The publications identified in this review present a relatively broad range of accuracy results of HPV testing on self-collected samples. However, this is partially explained by heterogeneity in the study methodology and it is concluded that, while the accuracy of testing on self-collected samples in lower than on clinician-collected samples, it is not substantially lower and can be considered a suitable alternative for women who do not attend for primary screening. An investigation into the most accurate HPV testing methods may be of use to optimise the accuracy of testing on self-collected samples and further validate the use of this method in a screening programme.

Summary: Criterion 4 **met** for self-sampling

2.2.1.2 Question 5 – Does self-collection of vaginal specimens increase uptake of cervical screening?

Criterion 6 of the UK NSC Screening Criteria states that: 'The test, from sample collection to delivery of results, should be acceptable to the target population.'

This review looked for studies reporting the uptake of self-sampling as a screening method for HPV in populations of under-screened women.

2.2.1.2.1 Description of the evidence

This review identified one SLR and meta-analysis,[38] and 10 additional primary studies,[39-48] all reporting self-sampling participation in 'under-screened' populations.

The SLR, Verdoodt 2015, identified 16 studies and evaluated whether offering a self-sampling kit could increase screening attendance in irregularly-screened or never-screened women, or women who did not respond to ≥1 invitation for conventional screening.[38] These populations align with the eligibility criteria in this review. Further criteria applied in Verdoodt 2015 were the exclusion of studies without a comparator arm, and studies with less than 1,000 participants;[38] as a result, the SLR would have excluded some studies which would have been eligible for inclusion in this review. However, these would have been small, non-comparative studies which would be unlikely to change the overall weight of the evidence.

2.2.1.2.2 Quality assessment

An assessment of the methodological quality of the SLR, using the AMSTAR checklist, demonstrated moderate quality, with 6 of 11 checklist questions addressed.[38] The points which were not addressed were mostly in relation to reporting and were not considered to be of great concern to the overall SLR quality.

Verdoodt 2015 assessed the quality of its included studies using the Cochrane tool for bias and reported overall moderate to high study quality (Table 4).[38] The studies were all conducted in 'under-screened' populations, which is the population of interest in this review question. The eligibility criteria with regards to age of participants in the studies were generally well aligned to the UK screening population, ranging from a minimum of 25 to 39 years old to a maximum of 50 to 69 years old. Of note in this quality assessment is that the exact time interval which was set as a threshold for 'participation' was not reported in a quarter of studies. A high risk of bias for 'selective reporting' was assigned to one study because women who had undertaken conventional screening in the clinic were removed entirely from the total number of women in the self-sampling arm, as opposed to presenting a per-protocol result (that is, the uptake of self-collected sampling amongst the entire population who were offered self-sampling). Women who opted out from the study were also removed entirely from analyses. Both omissions have potential to bias the result.

Table 4 Quality assessment (Cochrane tool for bias) of the studies included in Verdoodt 2015[38]

Risk of bias	Selection		Attrition	Reporting	
	Random sequence generation	Allocation concealment	Incomplete outcome data	Reporting of timelines	Selective reporting
Low	9	6	16	12	12
Moderate	7	10	0	4	3
High	0	0	0	0	1

The 10 additional primary studies identified in this review were assessed for quality using a modified Downs and Black checklist, the results from which are presented in Table 5.[39-48]

Table 5 Quality assessment (modified Downs and Black) of the primary studies included in this review

Question	Bosgraaf 2015[39]	Ducancelle 2015[40]	Duke 2015[41]	Enerly 2016[42]	Sultana 2016[43]	Tamalet 2016[44]	Verhoef 2014[45]	Virtanen 2015[46]	Kitchener 2016[47]	Racey 2016[48]
EXTERNAL VALIDITY										
Summary	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
CONFOUNDING										
Summary	Low	Low	High	Moderate	Low	Low	Moderate	Low	Moderate	High
POWER										
Summary	High	N/A	Unclear	Unclear	Unclear	N/A	N/A	N/A	Low	Unclear

Across all 10 identified studies, a low risk of bias is associated with 'external validity'.[39-48] To a certain extent, all studies assessed the population of interest to this review (women eligible for cervical cancer services but who do not participate in clinician led screening), although some were slightly less applicable, for example, Duke 2015 recruited from a generally 'under-screened' population instead of identifying individual women who had not responded to a screening invitation.[41] An additional concern is that none of the study populations were in age

ranges which align with UK cervical cancer screening eligibility,[39-48] however, this is unlikely to have a major impact on outcomes. Overall the studies are considered applicable.

The primary studies had an overall moderate risk of bias with respect to confounding. All studies included an accurate measure for participation outcome, but in 3 studies the number of samples returned but which had inadequate sample for testing were not reported.[41, 45, 47] The risk of bias relating to confounding was generally unclear as a result of under-reporting of baseline characteristics for the study populations. Among comparative studies, only 2 studies reported prognostic factors for participants: Sultana 2016 reported participant age and Racey 2016 additionally reported lifetime smoking history and number of sexual partners.[43, 48] As a result of the majority of studies not reporting baseline characteristics, the impact of these factors, or the requirement for adjustment to analyses to account for these, was unclear.

Many of the studies were non-comparative. For 4 comparative studies, power calculations were not reported,[41, 42, 48] or not reported for the outcome of interest.[43] One study reported power calculations but did not meet its prespecified required sample size: 16,500 women in each arm were required to give 80% power to detect a 1.4% difference in participation rates, however only 30,130 women were randomised.[39] Only a single study reported that it had adequate power to detect a meaningful difference between study arms.[47]

2.2.1.2.3 Results

Participation in self-sampling and clinician-collected sampling was reported in Verdoodt 2015 separately for studies utilising an intention to treat (ITT) or per-protocol analysis approach.[38]

The distinction between intention to treat (ITT) and per-protocol analyses is as follows:

- ITT results consider all participants who are screened including both those participating in self-sampling and those who attend for clinician-collected samples
- per-protocol analyses only consider the number of participants returning self-samples

Pooled analysis results from Verdoodt 2015 are presented in Table 6.[38]

Table 6 Verdoodt 2015 absolute participation results in ITT and per-protocol analyses[38]

		Absolute participation		Relative participation (95% CI)	Participation difference, % (95% CI)
	Studies, n	Self-sampling, % (95% CI)	Clinician-sampling, % (95% CI)		
Per-protocol					
Mail-to-all	13	20.7 (16.9 to 24.8) [range: 6.4 to 34.0]	10.3 (6.2 to 15.2)	2.06 (1.44 to 2.96)	9.9 (5.8 to 13.9)
Opt-in	3	9.7 (6.5 to 13.5)	12.2 (10.9 to 13.6)	0.72 (0.53 to 0.99)	-3.2 (-6.6 to 0.1)
ITT					
Mail-to-all	13	23.6 (20.2 to 27.3) [range: 10.2 to 39.0]	10.3 (6.2 to 15.2)	2.40 (1.73 to 3.33)	12.6 (9.3 to 15.9)
Opt-in	3	14.0 (8.0 to 21.4)	12.2 (10.9 to 13.6)	0.97 (0.65 to 1.46)	0.2 (-4.5 to 4.9)

ITT, intention to treat

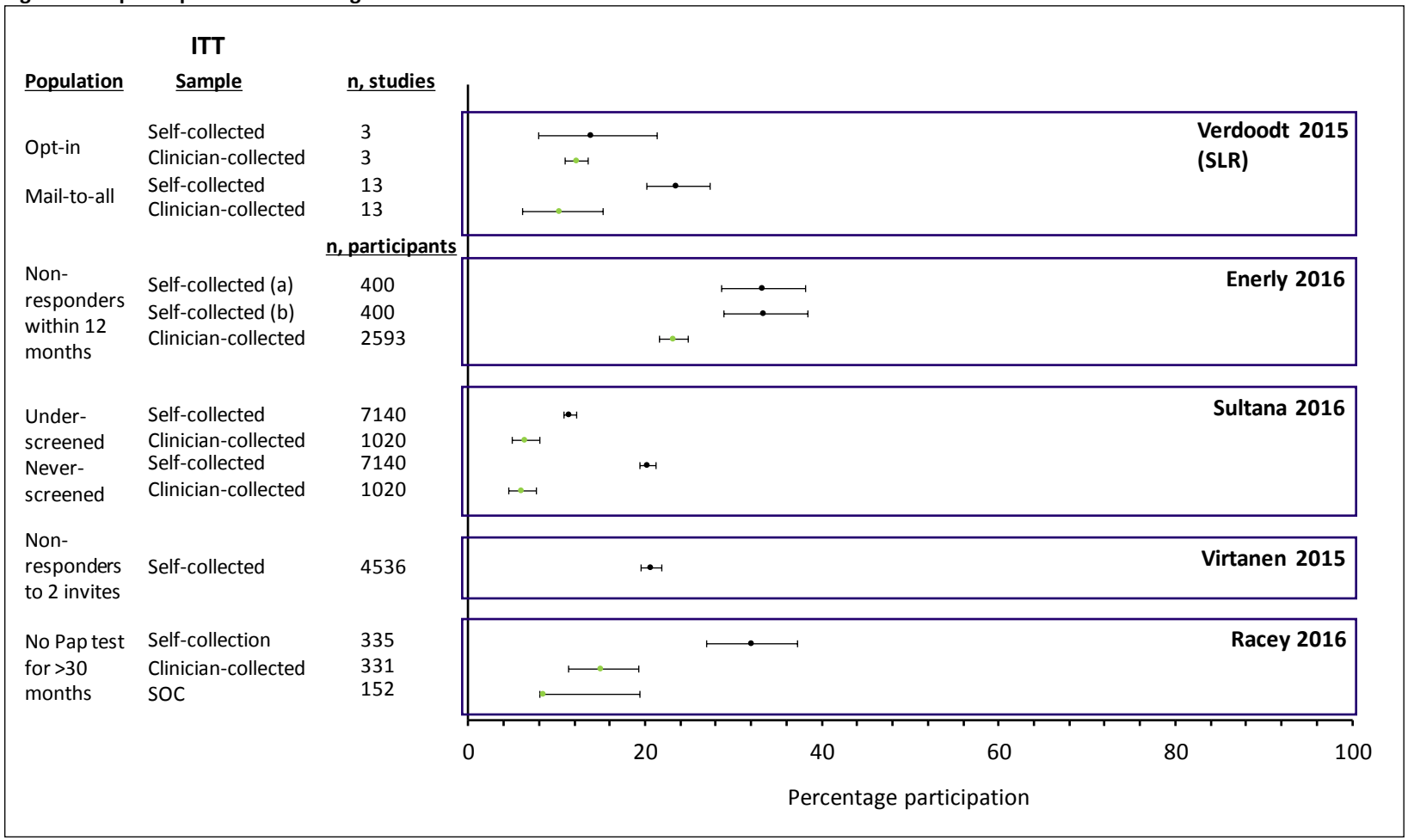
Verdoodt 2015 pools results separately for 'mail-to-all' or 'opt-in' study designs because the participation between these different distribution scenarios varied significantly. In mail-to-all studies, all participants were sent self-sampling kits directly to their home addresses, whereas in

opt-in study designs, women were sent an invitation to order a self-sampling kit by phone or mail, or alternatively to pick up a kit at a pharmacy. In mail-to-all studies, both per-protocol and ITT participation was significantly higher in the self-sampling arm than in the control arm. However, the results pooled across studies with an opt-in design did not demonstrate a significant difference between the self-collected and clinician-collected arms in either per-protocol or ITT analyses.[38] The results here highlight the impact of self-sample kit distribution on screening participation. The comparative ITT and per-protocol uptakes are discussed below.

Intention to treat participation

Four primary studies,[42, 43, 46, 48] in addition to the pooled analyses in Verdoodt 2015,[38] reported an ITT analysis. Results from these studies are presented in Figure 7.

Figure 7 ITT participation in screening



ITT, Intention to Treat; SOC, Standard of Care; SOC defined as opportunistic screening.

Self-collected (a): Evalyn Brush sampling device; Self collected (b): Delphi Screener sampling device

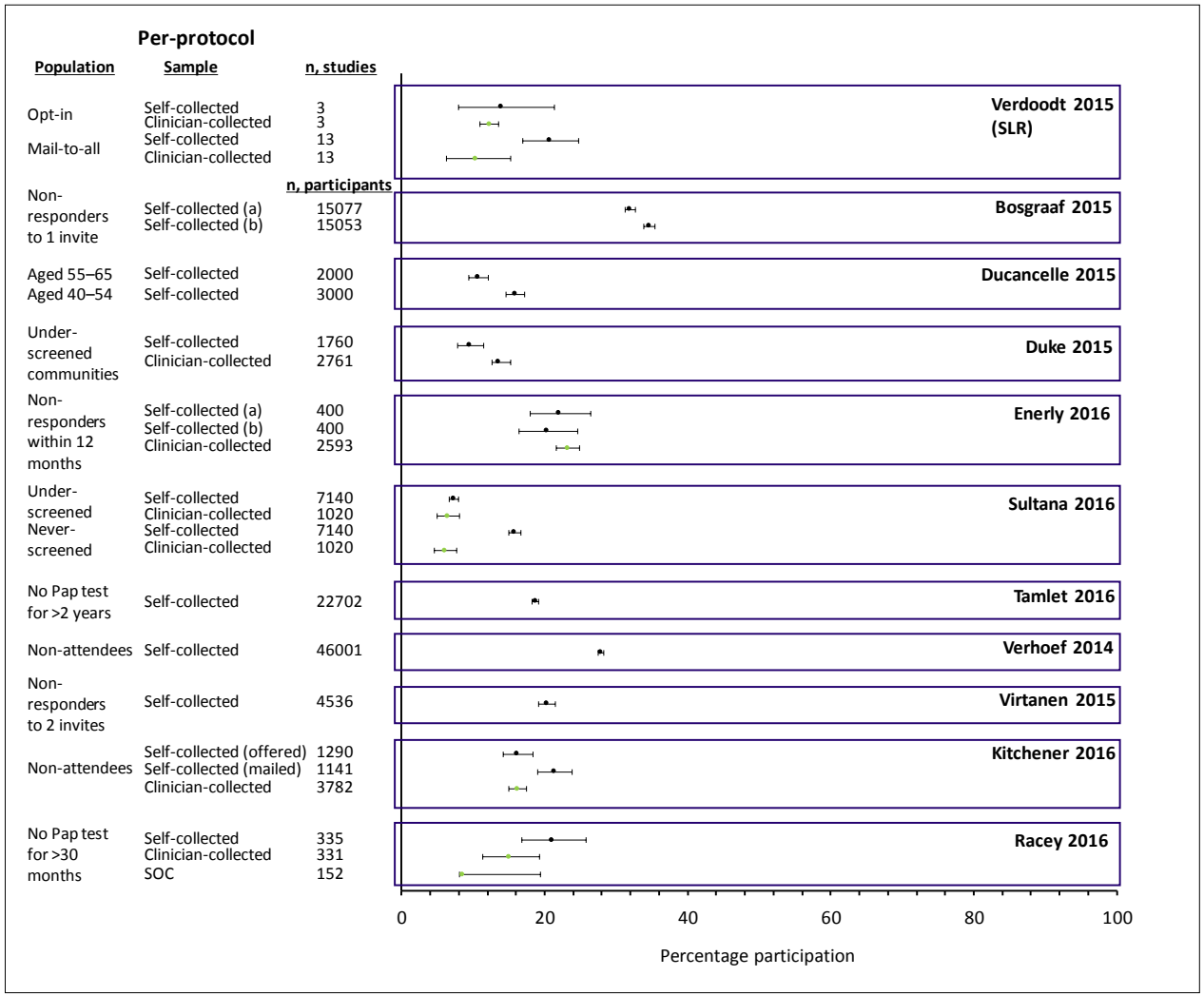
In all studies which reported participation in a control arm, the participation in screening was higher for participants offered self-collected sampling than for those who were offered clinician-sampling. Verdoodt 2015 reported an ITT uptake pooled across the mail-to-all studies of 23.6%, whereas in the comparable clinician-collected arm the uptake was 10.3%.[38] Absolute ITT participation in women offered self-collected sampling in the additional primary studies ranged from 11.5%[43] to 33.5%,[42] demonstrating a general agreement with the meta-analysis results.

A generally consistent result for screening uptake was reported across a well conducted SLR and 4 additional publications with no key quality concerns.[38, 42, 43, 46, 48] As a result, the results presented above are likely to be reliable.

Per-protocol participation

All 10 included primary studies,[39-48] in addition to Verdoodt 2015,[38] reported a per-protocol analysis. Results of these studies are presented in Figure 8.

Figure 8 Per-protocol participation in self-sampling



SOC: Standard of Care; SOC defined as opportunistic screening. Self-collected (a): Evalyn Brush sampling device; Self collected (b): Delphi Screener sampling device

Per-protocol participation in the self-collection arm, defined as the proportion of women returning self-sampling kits, was higher in a number of studies when compared to the clinician-collected participation, however, this was not always the case. The trend is not as apparent as in ITT analyses.

Verdoodt 2015 reported a per-protocol participation in the self-collection arm, pooled across mail-to-all studies, of 20.7%, whereas uptake in the clinician-collected arm was 10.3%. On consideration of opt-in study designs, the opposite trend was demonstrated and 9.7% of women in the self-collection arm returned samples, whereas participation in the clinician-collected arm was 12.2%.[38] Results from the self-collection arms of the 10 additional studies considering per-protocol participation ranged from 7.3%[43] to 34.6%.[39] These values are generally consistent. On consideration of the moderate to good quality of the primary studies, and particularly the low risk of bias associated with external validity, the results from the primary studies support the accuracy of the values reported in the meta-analysis.

As discussed previously, the SLR demonstrated an impact of self-sample kit distribution method on the uptake of self-collected sampling.[38] This outcome is further supported in the results presented here. Duke 2015 reports a per-protocol self-collected sampling uptake of 9.5% which is low in the overall range presented.[41] This result can be rationalised by the 'opt-in' methodology of the study and aligns closely with the 'opt-in' result from the meta-analysis. Furthermore, Kitchener 2016 assesses the difference in participation if self-sampling kits are sent to participants (analogous to a mail-to-all study design) or if kits are offered (opt-in design). Participation results are 21.3% and 16.2% respectively which adds further confidence in the SLR results.[47]

Intention to treat vs per-protocol

On consideration of the ITT and per-protocol participation results in Verdoodt 2015,[38] it is clear that sending a self-sampling kit can act as a prompt to encourage women to be screened and that some subsequently choose to attend clinician-sampling as opposed to returning a self-collected sample. In Verdoodt 2015, 20.7% of women returned self-sampling kits, whereas 23.6% of women attended any form of screening for cervical cancer.[38]

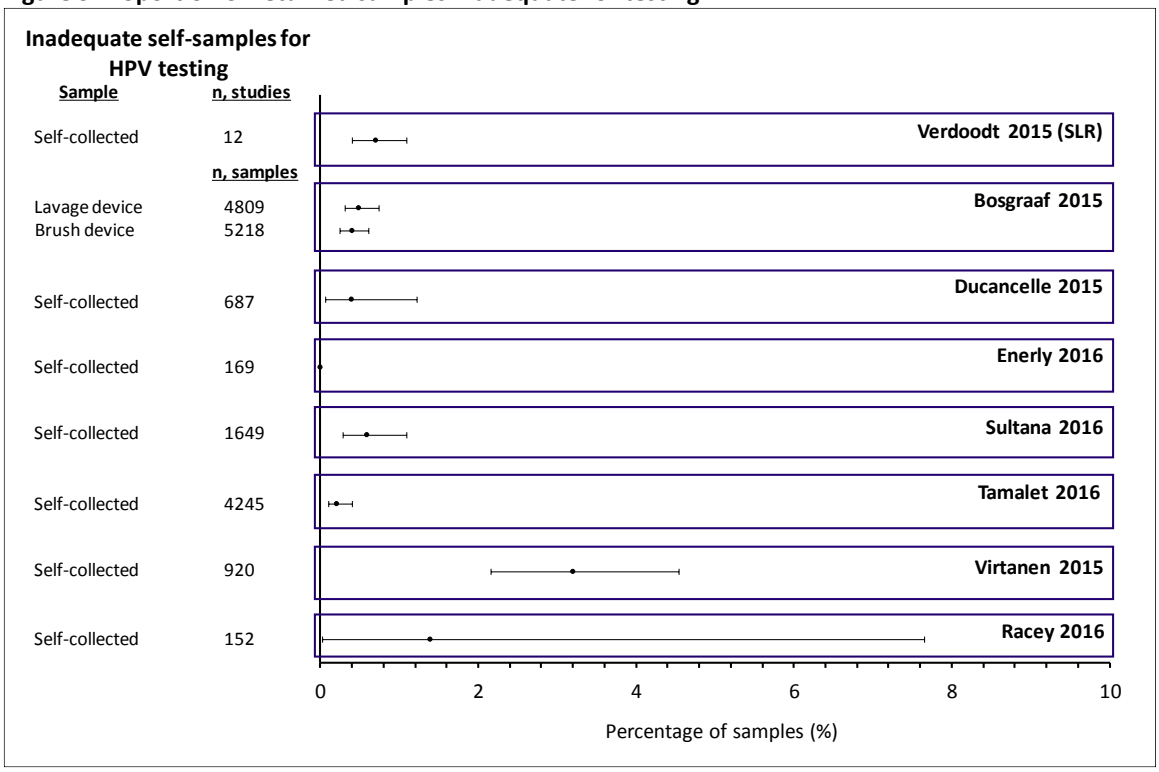
This trend was demonstrated to a greater extent in 3 of the additional primary studies reporting an ITT analysis.[42, 43, 48] Most notably, Enerly 2016 reported a screening participation in women offered a self-sampling kit of 33.4%; this value comprised 21.1% of women returning a self-collected sample and 12.3% attending for cytology at a clinic.[42] Similar results were seen in the other 2 studies;[43, 48]among the women who participated in screening after self-sampling was offered, approximately one-quarter to one-third chose instead to have clinician-collected sampling . In contrast, Virtanen 2015 reported a screening participation rate of 20.7%, of which 98% returned a self-collected sample.[46] This inconsistency between study results leads to uncertainty in the proportion of women who are offered self-sampling but instead attend for clinician-collected sampling.

An investigation into the potential impact this effect could have on the cost-effectiveness of a screening strategy would be beneficial. While this is a positive outcome with regards to increasing screening coverage, there is potential for an economic impact if a high proportion of women leave self-sampling kits unused and instead choose to visit a clinician. Additionally, it should be considered whether alternative approaches could increase screening uptake in a more cost-effective way.

Inadequate samples

The proportion of self-collected samples returned which were inadequate for HPV testing was presented in Verdoodt 2015,[38] pooled across 12 of its included studies, and was additionally presented in 8 of the primary studies identified in this review.[39, 40, 42-44, 46, 48] The proportions of inadequate samples from each study is presented in Figure 9.

Figure 9 Proportion of returned samples inadequate for testing



The pooled proportion of unsatisfactory samples reported in Verdoodt 2015 was 0.7% (95% CI 0.4% to 1.1%).[38] Similar results were presented in the additional primary studies.[39, 40, 42-44, 46, 48] The proportion of inadequate self-collected samples returned was 0[42] to 3.2%[46] across the studies, with the upper end of the 95% CI reaching 7.66% in one small study.[48] This supports the evidence in the SLR that concern associated with insufficient samples for HPV testing is low.[38]

2.2.1.2.4 Evidence summary

The evidence identified for this review question comprised one SLR and meta-analysis,[38] and 10 further primary studies reporting the uptake of self-collected sampling for HPV testing.[39-48] These publications have been demonstrated to be of moderate to good quality, and support the conclusion that offering previously under-screened women the option of self-collected sampling leads to a moderate increase in the overall screening uptake in this population. The increase in screening uptake is less pronounced when considered as an absolute difference to women in the clinician-collected arms of studies. However, since these studies are undertaken in populations of women who have previously not attended for primary screening, even a low increase in uptake could be considered successful, if a cost-effective strategy can be determined. Further investigation would be beneficial to determine if this is feasible. The evidence also suggests that the use of self-collection devices to collect samples for HPV sampling can be considered adequate, with only a very low proportion of insufficient samples reported.[38-48]

The SLR results demonstrate that participation in self-sampling is higher when sampling kits are mailed directly to eligible women, as opposed to inviting women to order or collect a kit.[38] This conclusion is supported further in primary studies and the method of distribution of self-sampling kits should be a key consideration in the design of a self-collected screening strategy.[41, 47] However, distributing kits to all eligible women when only approximately 1 in 5 will be returned could be costly and an assessment of the economic impact of this methodology could help ensure this is the most cost-effective method to increase uptake of screening.

Evidence presented here demonstrates that ITT analyses show generally improved participation rates when compared to per-protocol analyses.[38, 42, 43, 46, 48] This result could also affect the cost-effectiveness of a self-collection screening programme due to the cost of distributing self-sampling kits which remain unused, in addition to the higher cost of women subsequently attending for clinician-collected sampling, and would also warrant further investigation.

A final key consideration relating to this review question is the potential impact on primary screening uptake if women are aware that non-attendance will result in being sent a self-sampling kit, which might be considered more convenient than attending for clinician-based screening. Ensuring that implementation of a self-collection screening strategy is effective in reaching under-screened women, without encouraging women who would normally attend screening to miss clinician appointments, would be of great importance. An assessment of the impact that a subsequent 'self-sampling opportunity' would have on initial screening uptake would be of value.

Analysis of evidence relevant to criterion 6 – Uptake of self-collected sampling

Quantity: One SLR and meta-analysis[38] evaluating participation in self-collected sampling is supported by 10 primary studies identified in this review.[39-48] The SLR reported results from 16 studies, considering a total of 163, 431 women.[38] The evidence base for this review question is therefore large.

Quality: The quality of the SLR was assessed to be moderate using the AMSTAR checklist, addressing 6 of the 11 checklist questions.[38] The SLR reported moderate to high quality of its identified studies and highlighted no key quality concerns relating to this review.[38] An assessment of the 10 primary studies identified in this review demonstrated an overall moderate to good quality, particularly with regards to external validity for which all publications had a low risk of bias.

It should be noted that in 3 studies the proportion of samples returned which were inadequate for HPV-testing was not reported.[41, 45, 47] Additionally, the risk of bias associated with confounding was unclear because the majority of studies did not report population prognostic factors. However, it is unlikely that this would bias results significantly when considering uptake.

Applicability: The studies included in the SLR investigated populations of women considered to be 'under-screened' which aligns with the population of interest for this review question. Verdoodt 2015 only identified studies for inclusion which had a comparator arm, which was not an eligibility criterion in this review and could potentially impact the applicability of the SLR result.[38] None of the primary studies identified in this review assessed populations in the exact age range applicable to the UK cervical cancer screening population; in most instances this is unlikely to bias results, however, some studies were conducted in only younger[47] or older populations.[40]

Consistency: Generally consistent results were reported for self-sampling participation in both ITT[38, 42, 43, 46, 48] and per-protocol analyses.[38-48] This strengthens the conclusion that offering self-collected sampling to previously under-screened women could increase overall screening uptake. Additionally, a consistently low proportion of samples which were insufficient for HPV testing was reported across the studies,[38-40, 42-44, 46, 48] demonstrating that this is an adequate collection method.

Conclusions

A generally consistent result for participation was reported in the included studies, for both ITT and per-protocol analyses. This demonstrates that offering previously under-screened women the option of self-collected HPV testing leads to a moderate increase in the overall uptake in screening. However, while this is not insubstantial, the increase in uptake when compared to women sent another invitation for clinician sampling in the same studies was relatively small. Overall this result would benefit from further exploration.

The results also demonstrate that the use of a self-collection device is an adequate method of sample collection, with only very low proportions of samples reported to be inadequate for HPV testing.

The potential impact of offering self-sampling on the uptake of more accurate, clinician-based screening should be assessed to ensure a self-sampling strategy would not impact negatively on the overall detection of HPV. Additionally, an investigation into the circumstances in which the strategy would be useful and the most appropriate method of self-sample kit distribution, with regards to optimising uptake and cost-effectiveness, would aid the design of a self-sampling screening strategy.

Summary: Criterion 6 **uncertain** for self-sampling

3 Review summary

3.1 Conclusions and implications for policy

Summary of findings

Overall, this review found that:

- There is good evidence that self-collected sampling for HPV testing has lower accuracy than clinician tested sampling, but the difference is small
- There is good evidence that offering self-sampling could moderately increase the uptake of screening among women who do not respond to invitations for clinician-based screening, however, further investigation into the optimisation of a self-sampling strategy is required

3.1.1 Use of self-sampling to improve cervical screening uptake

The evidence identified in this review considering the use of self-collected samples for HPV screening demonstrated that offering previously under-screened women the option of self-collected sampling moderately increases overall screening uptake in this population. Additionally, it has been demonstrated that the use of a self-collection device is an adequate method for sample collection and that testing on these samples has sufficient accuracy to be considered an appropriate alternative to clinician-collected sampling where required. The evidence from which these conclusions are drawn is of moderate to good quality and with a reasonable consistency in results. However, the practical application of a self-collection strategy may need further exploration to ensure successful implementation and cost-effectiveness.

It should be assessed whether awareness that self-sampling would subsequently be available if primary screening is missed reduces uptake of standard clinician-based screening. All of the studies identified in this review investigated women who had failed to respond to 1, or a number of, previous screening invitations without being aware of a later self-sampling option and, as a result, did not investigate the impact this could have on initial screening uptake. Given that the accuracy of testing on self-collected samples is lower than on clinician-collected samples, it is important that women are not discouraged from attending clinician-sampling in favour of self-sampling. A Dutch model assessed the impact of women switching from clinician-collected to self-collected screening on the cost-effectiveness of a screening programme. The model varied the accuracy of the self-sampling test, the increase in attendance after offering self-sampling and risk of women developing cervical cancer. Under all scenarios, switching of women from clinician-collected to self-sampling resulted in a decrease in QALYs gained. If self-sampling were to be implemented it would be important to communicate the relative advantages of clinician-collected screening to reduce the impact of switching.[52]

The circumstances in which the strategy should be used is an important consideration. An additional consideration which warrants further investigation is the method of distributing self-sampling kits, and the impact this has on both the uptake and cost-effectiveness of screening. Verdoodt 2015 demonstrated that participation in self-sampling is higher when sampling kits are mailed directly to eligible women, as opposed to inviting women to order or collect a kit.[38] However, while sending all under-screened women self-sampling kits may increase screening uptake, distributing kits to all eligible women when only a small proportion are returned could

have a sizeable economic impact on a screening programme. An assessment of the threshold return rate at which this methodology would be cost-effective should be made to help ensure that this is a cost-effective method to increase screening participation. Modelling could be proposed to evaluate this factor.

3.2 Limitations of this review

This rapid review was conducted in line with the UK NSC requirements for evidence summaries.[53] These requirements are mostly in line with published guidelines for systematic reviews, but allowing for some methodological compromises. Some specific limitations relating to this review are discussed below.

3.2.1 Included study designs

It should be noted that this review was only designed to look for primary evidence that directly addressed the questions being considered. This review did not aim to systematically identify alternative forms of evidence such as modelling studies. A more holistic view of the evidence, such as modelling using published data on test accuracy and natural history, might be sufficient to give confidence that longer screening intervals would be safe.

3.2.2 Included publication types

This review only included peer-reviewed journal publications, and excluded any literature that was not peer-reviewed such as congress presentations and government reports. This may have led to the exclusion of relevant evidence that has only been published in non-peer-reviewed formats. However, this is an accepted methodological adjustment for a rapid review, and is unlikely to miss any pivotal studies, which would likely be published in peer-reviewed journals.

3.2.3 Review methodology

Articles were reviewed by a single reviewer in the first instance. A second reviewer examined all included articles, 10% of excluded articles, and any articles where there was uncertainty about inclusion. Although a fully systematic review would require all articles to be reviewed by both reviewers, this pragmatic strategy should have ensured that any articles where the eligibility was unclear were reviewed twice.

4 Appendices

Appendix 1 – Search strategy

Electronic databases

The search strategy included searches of the databases shown in **Error! Reference source not found.** and Table 7. MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase were searched simultaneously.

Table 7 Summary of electronic database searches and dates - use of self-sampling to improve screening uptake

Database	Platform	Searched on date	Date range of search
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid SP	19th October 2016	1974 to 2016 October 19
Embase	Ovid SP	19th October 2016	1974 to 2016 October 19
The Cochrane Library, including: • Cochrane Database of Systematic Reviews (CDSR) • Cochrane Central Register of Controlled Trials (CENTRAL) • Database of Abstracts of Reviews of Effects (DARE)	Wiley Online	20th October 2016	CENTRAL: Issue 9 of 12, September 2016 DARE: Issue 2 of 4, April 2015

Search terms

Search terms included combinations of free text and subject headings. Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase are shown in **Error! Reference source not found.** (Topic 1) and Table 8 (Topic 2), and search terms for the Cochrane Library databases are shown in **Error! Reference source not found.** (Topic 1) and Table 9 (Topic 2).

Table 8 Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase (searched simultaneously via Ovid SP) - use of self-sampling to improve screening uptake

Term group	#	Search terms	Results
Cervical cancer	1	exp Uterine cervical neoplasms/ or exp Uterine cervix cancer/ or exp Cervical intraepithelial neoplasia/ or uterine cervix carcinoma in situ/	166564
	2	(cervi$ adj3 (cancer$ or tumor$ or tumour$ or neoplas$ or dysplas$ or carcinoma$ or malignan$ or adenocarcinoma$ or choriocarcinoma$ or orteratoma$ or sarcoma$ or precancer$ or pre-cancer$)).tw.	151796
	3	(CIN or CIN1$ or CIN2$ or CIN3$ or "CIN 1$" or "CIN 2$" or "CIN 3$" or CINI$ or "CIN I$").tw.	23959
	4	or/1-3	213824
HPV	5	exp papillomavirus infections/ or exp papillomaviridae/ or exp human papilloma virus/ or exp papovavirus/	116054
	6	(human papillomavir$ or human papilloma vir$ or HPV$ or HR-HPV$ or hrHPV$).tw.	93535
	7	5 or 6	142016
Screening	8	exp mass screening/	310262

Term group	#	Search terms	Results
	9	exp Vaginal smears/ or human papillomavirus DNA tests/ or DNA probes, HPV/	58611
	10	(test$ or cotest$ or co-test$ or screen$).ti.	1074508
	11	case find$.tw.	9207
	12	((human papillomavir$ or human papilloma vir$ or HPV$ or HR-HPV$ or hrHPV$) adj2 (DNA or test$)).tw.	25002
	13	((vagina$ or cervi*) adj3 (smear$ or swab$ or scrap$ or test$ or sampl$)).tw.	37904
	14	exp Early detection of cancer/ or exp *Uterine cervical neoplasms/pc or exp *Uterine cervix cancer/pc or exp *Cervical intraepithelial neoplasia/pc or exp *Uterine cervix carcinoma in situ/pc	25327
	15	(annual adj (surveillance or review)).tw.	2464
	16	or/8-15	1350922
Self-collection	17	(self sampl$ or self collect$ or self care$ or self test$).tw.	34770
Outcome 1: Diagnostic test accuracy	18	exp "Sensitivity and Specificity"/ or *Diagnostic Accuracy/	763286
	19	(sensitiv$ or specific$ or accura$).tw.	8275328
	20	((pre-test or pretest) adj probability).tw.	4488
	21	((post-test or posttest) adj probability).tw.	1544
	22	predictive value$.tw.	204444
	23	likelihood ratio$.tw.	26869
	24	or/18-23	8647872
	25	limit 24 to yr=2013-2016	2056259
Outcome 2: Uptake/compliance	26	exp Patient acceptance of health care/ or exp Patient Attitude/	515638
	27	(satisf$ or dropout$ or drop out).tw.	624360
	28	(compliance or complie$ or comply$).tw.	263153
	29	(encourage$ or improve$ or improving or increas$ or promot$).tw.	14094754
	30	(uptake or particip$ or nonattend$ or non-attend$).tw.	2590532
	31	(accept$ or attend$ or attitude$ or utilisation or utilization).tw.	1696361
	32	(refus$ or respon$ or reluctan$ or nonrespon$).tw.	6484726
	33	or/26-32	20042442
	34	limit 33 to yr=2015-2016	2421932
Outcomes	35	25 or 34	3881443
Total	36	4 and 7 and 16 and 17 and 35	348
	37	exp animals/ not exp humans/	8792525
	38	("Journal: Conference Abstract" or comment or letter or case reports).pt.	6214134
	39	37 or 38	14681174
	40	36 not 39	301
	41	remove duplicates from 40	177

Table 9 Search terms for the Cochrane Library Databases (searched via the Wiley Online platform) - use of self-sampling to improve screening uptake

Term group	#	Search terms	Results
Cervical cancer	1	[mh "Uterine cervical neoplasms"] or [mh "Cervical intraepithelial neoplasia"]	1989
	2	(cervi* near/3 (cancer* or tumor* or tumour* or neoplas* or dysplas* or carcinoma* or malignan* or adenocarcinoma* or choriocarcinoma* or orteratoma* or sarcoma* or precancer* or pre-cancer*)):ti,ab,kw	3570
	3	(CIN or CIN1* or CIN2* or CIN3* or "CIN 1*" or "CIN 2*" or "CIN 3*" or CINI* or "CIN I*"):ti,ab,kw	1085
	4	{or #1-#3}	4198
HPV	5	[mh "papillomavirus infections"] or [mh papillomaviridae] or [mh "human papilloma virus"]	1218

Term group	#	Search terms	Results
	6	("human papillomavir*" or "human papilloma vir*" or HPV* or "HR-HPV*" or hrHPV*):ti,ab,kw	1593
	7	#5 or #6	1956
Screening	8	[mh "mass screening"]	5513
	9	[mh "Vaginal smears"] or [mh "test, hpv dna"]	798
	10	(test* or cotest* or co-test* or screen*):ti	23886
	11	"case find*":ti,ab,kw	215
	12	(("human papillomavir*" or "human papilloma vir*" or HPV* or "HR-HPV*" or hrHPV*) near/2 (DNA or test*)):ti,ab,kw	528
	13	((vagina* or cervi*) near/3 (smear* or swab* or scrap* or test* or sampl*)):ti,ab,kw	1650
	14	[mh "Early detection of cancer"] or [mh "Uterine cervical neoplasms"] or [mh "Cervical intraepithelial neoplasia"]	2727
	15	(annual next (surveillance or review)):ti,ab,kw	30
	16	{or #8-#15}	28112
Self-collection	17	("self sampl*" or "self collect*" or "self care*" or "self test*"):ti,ab,kw	5745
Outcome 1: Diagnostic test accuracy	18	[mh "Sensitivity and Specificity"] or [mh "Diagnostic Accuracy"]	18093
	19	(sensitiv* or specific* or accura*):ti,ab,kw	111264
	20	(("pre-test" or pretest) next probability):ti,ab,kw	86
	21	(("post-test" or posttest) next probability):ti,ab,kw	43
	22	"predictive value*":ti,ab,kw	11838
	23	"likelihood ratio*":ti,ab,kw	488
	24	{or #18-#24}	117737
	25	#24 Publication Year from 2013 to 2016	33032
Outcome 2: Uptake/compliance	26	[mh "Patient acceptance of health care"] or [mh "Patient Attitude"]	24278
	27	(satisf* or dropout* or "drop out*"):ti,ab,kw	37641
	28	(compliance or complie* or comply*):ti,ab,kw	26628
	29	(encourage* or improve* or improving or increas* or promot*):ti,ab,kw	381535
	30	(uptake or particip* or nonattend* or "non-attend*"):ti,ab,kw	128968
	31	(accept* or attend* or attitude* or utilisation or utilization):ti,ab,kw	59548
	32	(refus* or respon* or reluctan* or nonrespon* or "non-respon*"):ti,ab,kw	179564
	33	{or #26-#32}	535459
	34	#33 Publication Year from 2015 to 2016	53244
Outcomes	35	#25 or #34	75818
Total	36	#4 and #7 and #16 and #17 and #35 in Other Reviews and Trials	27

Appendix 2 – Study selection

Review process

The following review process was followed:

- Each abstract was reviewed against the inclusion/exclusion criteria by one reviewer. Where the applicability of the inclusion criteria was unclear, the article was included at this stage in order to ensure that all potentially relevant studies were captured. A second independent reviewer provided input in cases of uncertainty, and validated 20% of the first reviewer's screening decisions. Any disagreements were resolved by discussion until a consensus was met.
- Full-text articles required for the full-text review stage were acquired.
- Each full-text article was reviewed against the inclusion/exclusion criteria by one reviewer, who determined whether the article was relevant to one or more of the review questions. A second independent reviewer provided input in cases of uncertainty, and validated 20% of the first reviewer's screening decisions. Any disagreements were resolved by discussion until a consensus was met.

Eligibility criteria

Eligibility criteria for each question are presented in **Error! Reference source not found.** to Table 11 below. All search results for Topic 1 were reviewed against the eligibility criteria for Questions 1 to 3; all search results for Topic 2 were reviewed against the eligibility criteria for Questions 4 and 5.

For all topics, systematic reviews and meta-analyses were considered for inclusion in this review. If the scope of a systematic review or meta-analysis was very closely aligned to one of the questions in this review, it was included in this review in its own right. However, if the scope was not closely aligned to one of the questions in this review but some of the included articles were of interest, the reference list of the systematic review or meta-analysis was hand-searched. Any primary research articles that were identified as being relevant to this review were then included.

Table 10 Eligibility criteria for publications relating to question 4

Domain	Inclusion criteria	Exclusion criteria
Population	Women in the cervical cancer screening population	Studies that do not include women eligible for cervical cancer screening
Intervention (s)	HPV testing on a self-collected sample	Studies that do not include an HPV test on a self-collected sample
Comparator	HPV or cytology testing on a clinician-collected sample	Studies that do not include a comparator test on a clinician-collected sample
Reference Standard	Colposcopy or biopsy	Any other reference standard
Outcomes	Measures of screening accuracy, or sufficient data to calculate these: • Sensitivity • Specificity • Positive predictive value • Negative predictive value • Accuracy • Likelihood ratio	Outcomes not relating to the measures of screening accuracy of HPV testing on self-collected specimens
Study design and publication type	Peer-reviewed evidence derived from the following types of study: • RCTs	• Other study designs or publication types • Retrospective studies, case control studies or cross-sectional studies

Domain	Inclusion criteria	Exclusion criteria
	• Non-randomised, comparative interventional studies • Prospective cohort studies • Systematic reviews and meta-analyses of the above study types	• Conference abstracts or other publication types that have not been peer-reviewed
Language	English language	Non-English language

Table 11 Eligibility criteria for publications relating to question 5

Domain	Inclusion criteria	Exclusion criteria
Population	Women who are eligible for cervical cancer screening but do not participate in clinician-led screening services (note that these may be a sub-group of a larger study)	Studies that do not include women who are eligible for cervical cancer screening but do not participate in clinician-led screening services, or studies that do not report outcomes separately for this group
Intervention	Offer or invitation of HPV testing on a self-collected sample	Studies that do not include an offer or invitation of HPV self-sampling
Comparator	Offer or invitation of a clinician-collected sample, or no comparator	-
Outcomes	Measures of uptake, compliance or participation	Any other outcomes
Study design and publication type	Peer-reviewed evidence derived from the following types of study: • RCTs • Non-randomised, comparative interventional studies • Prospective cohort studies • Systematic reviews and meta-analyses of the above study types	• Other study designs or publication types • Retrospective studies, case control studies or cross-sectional studies • Conference abstracts or other publication types that have not been peer-reviewed
Language	English language	Non-English language

Appraisal for quality and risk of bias

The following tools were used to assess the quality and risk of bias of each study included in the review:

- **Systematic literature reviews:** Assessing the Methodological Quality of Systematic Reviews (AMSTAR) checklist[54]
- **Diagnostic accuracy studies**: Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[55]
- **Interventional and observational studies**: Modified versions of the Downs and Black checklist[56]

Reviewers were provided with guidance criteria to ensure consistent applicability of the quality assessment checklists for primary studies; these criteria are detailed in **Error! Reference source not found.** (Modified Downs & Black, Topic 1), Table 12 (QUADAS-2, Topic 2, Question 4) and Table 13 (Modified Downs & Black, Topic 2, Question 5).

Table 12 Template quality assessment checklist for question 4

Question	Literature-recommended criteria	Guideline criteria for HPV self-testing
PATIENT SELECTION		
Was a consecutive or random sample of patients enrolled?	A study should ideally enrol all consecutive, or a random sample of, eligible patients with suspected disease – otherwise there is potential for bias. Studies that make inappropriate exclusions, e.g. excluding "difficult to diagnose" patients, may result in overoptimistic estimates of diagnostic accuracy	**Yes** if all eligible women (asymptomatic women within the screening age range) or a random sample of women within the study period were included **No** if participants were selected in a different way, eg. by referral or convenience sample

Question	Literature-recommended criteria	Guideline criteria for HPV self-testing
Was a case-control design avoided?	Studies enrolling patients with known disease and a control group without the condition may exaggerate diagnostic accuracy	**Yes** if the study was a prospective cohort study **No** if cases of colposcopy-confirmed HPV were matched to controls with negative colposcopy results
Did the study avoid inappropriate exclusions?	Exclusion of patients with "red flags" for the target condition, who may be easier to diagnose, may lead to underestimation of diagnostic accuracy	**Yes** if all patients were included, or if exclusions were appropriate (such as women outside the low-risk screening population) and unlikely to lead to bias **No** if any group within the screening population was systematically excluded
Could the selection of patients have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the included patients do not match the review question?	There may be concerns regarding applicability if patients included in the study differ, compared to those targeted by the review question, in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols	**Low** if patients overall have a normal risk for HPV so are representative of the screening population; the population should be women who had not responded to invitations for clinical testing **High** if patients overall are not representative of the screening population
INDEX TESTS		
Were the index test results interpreted without knowledge of the reference standard?	This item is similar to "blinding" in intervention studies. Interpretation of index test results may be influenced by knowledge of the reference standard	**Yes** if self-testing results were interpreted before colposcopy was performed, or if the results were interpreted after colposcopy was performed but without knowledge of the colposcopy diagnosis **No** if the colposcopy diagnosis was known when self-testing results were interpreted
If a threshold was used, was it pre-specified?	Selecting the test threshold to optimise sensitivity and/or specificity may lead to overoptimistic estimates of test performance, which is likely to be poorer in an independent sample of patients in whom the same threshold is used	**Yes** when the threshold for a positive result on the HPV test was pre-specified, such as using the threshold specified by the manufacturer **No** when the threshold was not pre-specified
Could the conduct or interpretation of the index test have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Variations in test technology, execution, or interpretation may affect estimates of its diagnostic accuracy. If index tests methods vary from those specified in the review question there may be concerns regarding applicability	**Low** when the conduct and interpretation of the self-testing kits were relevant to the UK care setting **High** if any aspect of the index test, including its conduct or interpretation, was substantially different from the UK care setting
REFERENCE STANDARD		
Is the reference standard likely to correctly classify the test condition?	Estimates of test accuracy are based on the assumption that the reference standard is 100% sensitive and specific. Disagreements between the reference standard and index test are assumed to result from incorrect classification by the index test	**Yes** if the diagnosis was confirmed by colposcopy **No** if any other reference standard was used (note that HPV testing or cytology on a clinician-collected sample is a relevant *comparator* in this review but is not the *reference standard*)

Question	Literature-recommended criteria	Guideline criteria for HPV self-testing
Were the reference standard results interpreted without knowledge of the results of the index test?	Potential for bias is related to the potential influence of prior knowledge on the interpretation of the reference standard	**Yes** if colposcopy results were interpreted before self-testing was performed, or if the results were interpreted after self-testing was performed but without knowledge of the self-testing diagnosis **No** if the self-testing diagnosis was known when the colposcopy results were interpreted
Could the reference standard, its conduct, or its interpretation have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the target condition as defined by the reference standard does not match the review question?	The reference standard may be free of bias but the target condition that it defines may differ from the target condition specified in the review question. For example, when defining urinary tract infection, the reference standard is generally based on specimen culture but the threshold above which a result is considered positive may vary	**Low** if the target condition is cervical abnormalities (CIN2+, CIN3+, invasive cervical cancer) **High** for any other target condition
PATIENT FLOW		
Was there an appropriate interval between the index test(s) and the reference standard?	Ideally results of the index test and reference standard are collected on the same patients at the same time. If there is a delay or if treatment is started between index test and reference standard, misclassification may occur due to recovery or deterioration of the condition. The length of interval leading to a high risk of bias will vary between conditions. A delay of a few days may not be a problem for chronic conditions, while for acute infectious diseases a short delay may be important	**Yes** if the self-testing was conducted within a week of the clinical tests, on average **No** if the self-testing was conducted more than a week before or after the clinical tests, on average
Did all patients receive a reference standard? Did patients receive the same reference standard?	Verification bias occurs when not all of the study group receive confirmation of the diagnosis by the same reference standard. If the results of the index test influence the decision on whether to perform the reference standard or which reference standard is used, estimated diagnostic accuracy may be biased	**Yes, Yes** if all screened patients had confirmation of their diagnosis, and all were diagnosed in the same manner (similarly trained staff, similar timing of diagnosis) **No, Yes** if not all patients had colposcopy or biopsy, but those who did had the same reference standard **No** if patients received different reference standards
Were all patients included in the analysis?	All patients who were recruited into the study should be included in the analysis. There is a potential for bias if the number of patients enrolled differs from the number of patients included in the 2x2 table of results, for example because patients lost to follow-up differ systematically from those who remain	**Yes** if all screened patients were included in the final analysis **No** if any screened patients were not included in the final analysis
Could the patient flow have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain

Table 13 Template quality assessment checklist for question 5

Question	Guideline criteria for question 5
REPORTING	

Question	Guideline criteria for question 5
Is the hypothesis/aim/objective of the study clearly described?	Section removed – relates to reporting rather than study quality
Are the main outcomes to be measured clearly described in the Introduction or Methods section?	
Are the characteristics of the patients included in the study clearly described?	
Are the intervention(s) of interest clearly described?	
Are the distributions of principal confounders in each group of subjects to be compared clearly described?	
Are the main findings of the study clearly described?	
Does the study provide estimates of the random variability in the data for the main outcomes?	
Have all important adverse events that may be a consequence of the intervention been reported?	
Have the characteristics of patients lost to follow-up been described?	
EXTERNAL VALIDITY	
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Answer should relate to the outcome measures of interest (adherence, compliance, uptake)
Modified question: Were the subjects asked to participate in the study representative of the population of interest for this review? Original question: Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	**Yes** only when the target population was either all women eligible for screening, or women who had not responded to invitations for clinical testing **No** if study was performed in only a certain subgroup of the population of interest
Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?	Question removed – investigating new care settings, so not possible to be representative of the treatment the majority of patients currently receive
CONFOUNDING	
Was an attempt made to blind study subjects to the intervention they have received?	Question removed – not applicable to screening and surveillance methods
Was an attempt made to blind those measuring the main outcomes of the intervention?	Question removed – outcome is either returning a self-sampling kit (for intervention) or attendance at clinic (for comparator), so it would not be possible to blind the investigator measuring these outcomes
If any of the results of the study were based on "data dredging", was this made clear?	Question removed – only interested in adherence, compliance, uptake
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients? Removed wording: "…or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?"	**Yes** if analyses were adjusted for different lengths of follow-up if necessary, or if length of follow-up was comparable between groups **No** if the length of follow-up was not comparable, and analyses were not adjusted
Were the statistical tests used to assess the main outcomes appropriate?	**Yes** if groups were compared appropriately using risk difference, risk ratios, odds ratios, unpaired t-tests or similar; for single-arm trials a paired t-test may be appropriate; other methods may also be appropriate if justified in the publication **No** if the statistical tests were not appropriate – to be determined on a case-by-case basis

Question	Guideline criteria for question 5
	NA for single arm studies
Was compliance with the intervention/s reliable?	Question removed – compliance is an outcome of interest
Were the main outcome measures used accurate (valid and reliable)?	Answer should relate to the outcome measures of interest to this review (adherence, compliance, uptake) **Yes** when uptake was measured in a valid and reliable way, and the proportion of usable samples returned from self-testing was reported **Unclear** when uptake was measured in a valid and reliable way, but the proportion of useable samples has not been reported **No** if uptake was not measured in a valid and reliable way, or it is not clear how many samples returned from self-testing were usable
Were the patients in different intervention groups (trials and cohort studies) recruited from the same population? Removed wording: "…or were the cases and controls (case-control studies)…"	**Yes** if patients from all intervention groups were recruited from the same population **No** if different intervention groups were recruited from different populations, such as different geographical location or different baseline characteristics NA for single arm studies
Question added: Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	**Yes** if baseline characteristics were similar between treatment groups, particularly age and proportion with HPV vaccination **No** if there were significant differences between the groups in either of the characteristics listed above NA for single arm studies
Were study subjects in different intervention groups (trials and cohort studies) recruited over the same period of time? Removed wording: "…or were the cases and controls (case-control studies)…"	**Yes** if patients from all intervention groups were recruited over the same period of time **No** if patients from different intervention groups were recruited at different times, such as historical control groups NA for single arm studies
Were study subjects randomised to intervention groups?	**Yes** if randomisation was performed using computer-generated random numbers or random number tables **Inadequate** if alternation, case record numbers, birth dates or week days were used to allocate patients to treatment arms **No** if no attempt was made at randomisation
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	**Yes** if the allocation sequence was protected before and until assignment, using methods such as: centralised or pharmacy-controlled randomisation, serially-numbered identical containers, on-site computer-based system with a randomisation sequence that is not readable until allocation, or other approaches with robust methods to prevent foreknowledge of the allocation sequence **No** if inadequate methods of randomisation were used, or if random number lists could have been viewed before allocation, such as open random number lists or serially numbered envelopes NA in non-randomised studies
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Answer should relate to the outcome measures of interest to this review **Yes** if analyses were adjusted for differences in key baseline characteristics, or if adjustment was not necessary **No** if adjustment was necessary but was not performed NA for single arm studies
Were losses of patients to follow-up taken into account?	Question removed – loss to follow-up is related to compliance, which is an outcome of interest
POWER	

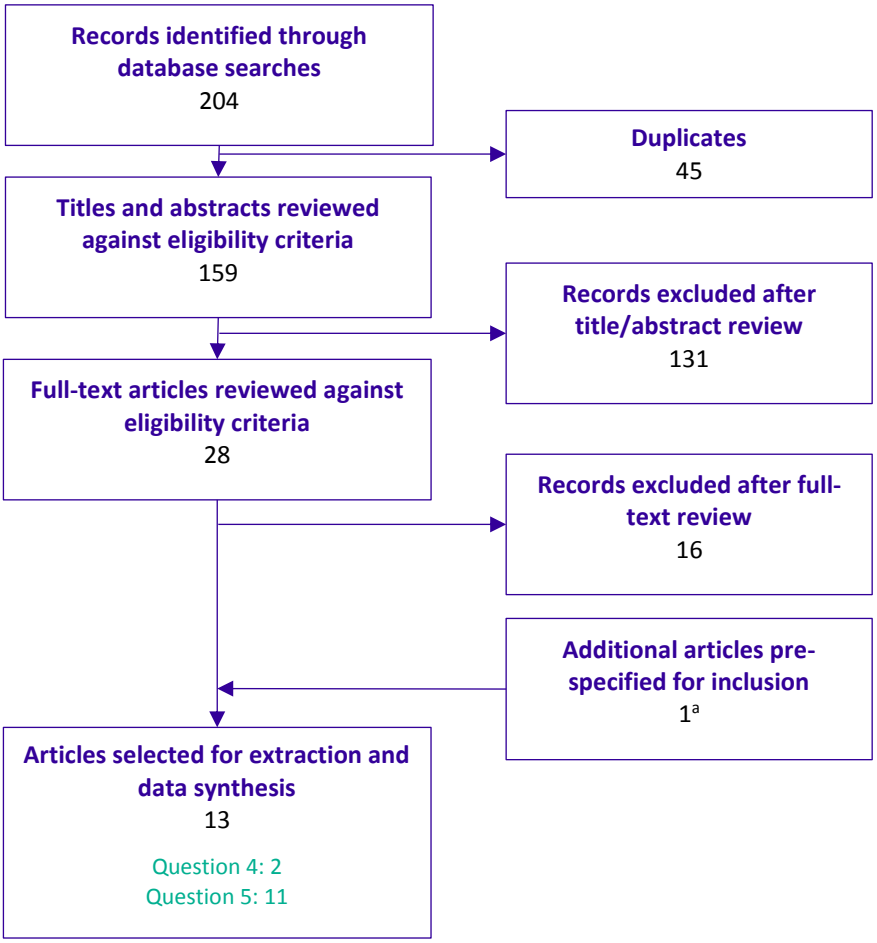
Question	Guideline criteria for question 5
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	**Yes** if power calculations are reported and an adequate sample size was used **No** if power calculations are reported and an adequate sample size was not reached **Unclear** if power calculations are not reported (adequate sample sizes may be calculated for each outcome when a clinically important difference has been determined) NA for single arm studies

Appendix 3 – Included and excluded studies

PRISMA flowcharts

Figure 13 summarises the volume of publications included and excluded at each stage of the review.

Figure 10 Summary of publications included and excluded at each stage of the review



[a] Verdoodt 2015 was identified as a relevant SLR at the protocol stage and was therefore included in the evidence synthesis

Appendix 4 – Study-level synthesis of results

Screening and surveillance intervals

Table 14 Quality assessments for the SLRs informing question 4 and question 5

Question	Arbyn 2014[35]	Verdoodt 2015[38]
Was an 'a priori' design provided?	Yes	Unclear
Was there duplicate study selection and data extraction?	Yes for selection, Unclear for data extraction	Yes
Was a comprehensive literature search performed?	Yes	Yes
Was the status of publication (ie. Grey literature) used as an inclusion criterion?	Yes	No
Was a list of studies (included and excluded) provided?	Yes	No
Were the characteristics of the included studies provided?	Yes	Yes
Was the scientific quality of the included studies assessed and documented?	Yes	Yes
Was the scientific quality of the included studies used appropriately in formulating conclusions?	Yes	No
Were the methods used to combine the findings of studies appropriate?	Yes	Yes
Was the likelihood of publication bias assessed?	Yes	No
Was the conflict of interest included?	No	No

Accuracy of self-sampling

Details of studies relevant to question 4 are presented in Table 15, and full quality assessments are presented in Table 16.

Table 15 Studies relevant to question 4

Study	Study design	Population characteristics	Screening methods	Epidemiology
Arbyn 2014[35] (SLR)	Design SLR and meta-analysis Objective To assess whether HPV testing on self-collected samples is equivalent to HPV testing on samples	Eligible studies Intervention: A vaginal sample was self-taken by a woman followed by a sample taken by a clinician or RCT with self-sample in 1 arm and clinician sample in the other A high-risk HPV DNA or RNA test was done on both samples or clinician sample examined microscopically for presence of	Self-testing In studies in a healthy screening population the devices used for self-sampling were: • Swab in 7 studies (5 papers) • Brush in 6 studies • Tampon in 1 study • Spatula in 1 study To define test positivity of the HPV test, the	Screening test accuracy Comparison of self-testing and clinician-testing: Pooled absolute sensitivity and specificity

	Self-sampling	Clinician-sampling		
		HPV	Cytology ASC-US+	Cytology LSIL+
CIN2 or worse				
Studies, n	16	16	12	8

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Epidemiology</th></tr>
<tr>
<td></td>
<td>collected by clinicians

<u>Dates</u>
Papers published between Jan 1, 1990 and June 3, 2013</td>
<td>cytological epithelial lesions

Presence or absence of CIN2+ verified by colposcopy or biopsy in all enrolled women with at least 1 positive test

<u>Patient recruitment:</u>
• Those attending routine cervical cancer screening (population of interest)
• High risk women
• Those referred to colposcopy because of previous positive screening results

<u>Data analysis</u>
The pooled absolute sensitivity and specificity of the tests were estimated jointly with metandi, a procedure in STATA, based on a bivariate model for the logit transformations of sensitivity and specificity, taking the intrinsic correlation between true positive and false positive rates and the variability between studies into account.

The relative sensitivity and specificity of HPV testing on self-samples compared with cytology HPV testing on clinician-taken samples using metadas, a SAS macro for the meta-analysis of diagnostic accuracy studies that allows the inclusion of type of test as a covariate, making comparison of tests possible.

<u>SLR results</u>
Sample size and demographics:

Data from 36 studies (reported in 34 papers) which altogether enrolled 154,556 women.

16 studies (14 papers) were in the population of interest for this review (primary screening of generally healthy</td>
<td>cut-off proposed by the manufacturer was accepted.

The reference standard (as per the inclusion criteria for the SLR) was colposcopy or biopsy in all enrolled women with at least 1 positive test.

<u>Clinician-testing</u>
34 of the selected studies used HPV testing on a clinician taken sample as the comparator. Additionally, the clinician taken samples were examined cytologically in 20 reports. 18 of these 20 studies included both cytology and HPV testing on the clinician-taken sample.

To define test positivity of the HPV test, the cutoff proposed by the manufacturer was accepted.

For cytological tests, 2 cutoffs were considered:

• Atypical squamous cells of undetermined significance (ASC-US) or worse
• Low-grade squamous intraepithelial lesions (LSIL) or worse
The reference standard (as per the inclusion criteria for the SLR) was colposcopy or biopsy in all enrolled women with at least 1 positive test.</td>
<td>
<table>
<tr><td>Sensitivity, % (95% CI) [range]</td><td>76 (69 to 82) [51% to 93%]</td><td>91 (87 to 94) [NR]</td><td>83 (75 to 89) [NR]</td><td>71 (66 to 76) [NR]</td></tr>
<tr><td>Specificity, % (95% CI) [range]</td><td>86 (83 to 89) [67% to 93%]</td><td>88 (85 to 91) [NR]</td><td>91 (87 to 94) [NR]</td><td>97 (97 to 98) [NR]</td></tr>
<tr><td colspan="5">CIN3 or worse</td></tr>
<tr><td>Studies, n</td><td>8</td><td>8</td><td>6</td><td>5</td></tr>
<tr><td>Sensitivity, % (95% CI) [range]</td><td>84 (72 to 92) [63% to 94%]</td><td>95 (91 to 97) [NR]</td><td>91 (85 to 95) [NR]</td><td>78 (72 to 85) [NR]</td></tr>
<tr><td>Specificity, % (95% CI)</td><td>87 (84 to 90)</td><td>89 (87 to 92)</td><td>89 (86 to 91)</td><td>97 (96 to 97)</td></tr>
</table>
Relative accuracy of HPV self-samples vs clinician-taken samples in all included studies^
<table>
<tr><th>CIN grade</th><th>Studies n</th><th>Relative sensitivity (95% CI)</th><th>Relative specificity (95% CI)</th></tr>
<tr><td colspan="4">HPV on self-samples vs HPV on clinician samples</td></tr>
<tr><td>CIN2 or worse</td><td>34</td><td>0.88 (0.85 to 0.91)*</td><td>0.96 (0.95 to 0.97)*</td></tr>
<tr><td>CIN3 or worse</td><td>12</td><td>0.89 (0.83 to 0.96)*</td><td>0.96 (0.93 to 0.99)*</td></tr>
<tr><td colspan="4">HPV on self-samples vs cytology (ASC-CU+) on clinician samples</td></tr>
<tr><td>CIN2 or worse</td><td>19</td><td>0.95 (0.91 to 0.99)*</td><td>0.92 (0.90 to 0.94)*</td></tr>
<tr><td>CIN3 or worse</td><td>6</td><td>0.99 (0.94 to 1.06)</td><td>0.98 (0.97 to 0.99)*</td></tr>
<tr><td colspan="4">HPV on self-samples vs cytology (LSIL+) on clinician samples</td></tr>
<tr><td>CIN2 or worse</td><td>11</td><td>1.14 (1.07 to 1.21)*</td><td>0.88 (0.86 to 0.90)*</td></tr>
<tr><td>CIN3 or worse</td><td>6</td><td>1.19 (1.09 to 1.29)*</td><td>0.90 (0.87 to 0.94)*</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Epidemiology</th></tr>
<tr>
<td></td>
<td></td>
<td>
women)

Most common exclusion criteria for women in studies were pregnancy (in 10 studies), hysterectomy (in 7 studies), prior pelvic radiation (7 studies) and recent history of screening (in 7 studies)

<u>Study quality assessments</u>
Methodological quality of all included studies was assessed by QUADAS checklist and was overall moderate to good.

<u>Risk of bias in all included studies</u>
<table>
<tr><th>Risk</th><th>Studies n (%), n=36</th></tr>
<tr><td colspan="2">Enrolment of patients</td></tr>
<tr><td>Low</td><td>19 (53%)</td></tr>
<tr><td>Moderate</td><td>16 (44%)</td></tr>
<tr><td>High</td><td>1 (3%)</td></tr>
<tr><td colspan="2">Reporting and execution of index and comparator test</td></tr>
<tr><td>Adequate</td><td>26 (72%)</td></tr>
<tr><td>Unclear</td><td>10 (28%)</td></tr>
<tr><td>High</td><td>0 (0)</td></tr>
<tr><td colspan="2">Quality of the verification with a reference standard</td></tr>
<tr><td>Good</td><td>32 (89%)</td></tr>
<tr><td>Moderate</td><td>3 (8%)</td></tr>
<tr><td>Possibly problematic</td><td>1 (3%)</td></tr>
<tr><td colspan="2">Delay between self-sampling, clinician sampling and verification with reference standard</td></tr>
<tr><td>Short (<6 months)</td><td>25 (69%)</td></tr>
<tr><td>Long</td><td>9 (25%)</td></tr>
<tr><td>Unreported</td><td>2 (6%)</td></tr>
<tr><td colspan="2">Partial verification</td></tr>
<tr><td>Avoided</td><td>28 (78%)</td></tr>
<tr><td>Present</td><td>8 (22%)</td></tr>
<tr><td colspan="2">Differential verification</td></tr>
</table>
</td>
<td></td>
<td>^Studies pooled across settings (healthy screening population, high-risk patients and patients at follow up) due to low variation in relative sensitivity and specificity between self-collected and clinician collected samples

*Statistically significant</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Epidemiology</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><td>Absent</td><td>35 (97%)</td></tr>
<tr><td>Present</td><td>1 (3%)</td></tr>
<tr><td colspan="2">Withdrawal of patients explained appropriately</td></tr>
<tr><td>Yes</td><td>25 (69%)</td></tr>
<tr><td>No</td><td>9 (25%)</td></tr>
<tr><td colspan="2">Poor reporting of uninterpretable assessed tests</td></tr>
<tr><td>Yes</td><td>20 (56%)</td></tr>
<tr><td>No</td><td>16 (44%)</td></tr>
<tr><td colspan="2">Poor reporting of uninterpretable reference standard</td></tr>
<tr><td>Yes</td><td>22 (61%)</td></tr>
<tr><td>No</td><td>12 (39%)</td></tr>
</table>
</td><td></td><td></td></tr>
<tr><td>Stanczuk 2016[36]

(PaVDaG study)</td><td><u>Design</u>
Prospective cohort study

<u>Objective</u>
To assess the performance of a high-risk human papillomavirus (hrHPV) PCR-based assay to detect CIN2+ in self-collected vaginal and urine samples

<u>Dates</u>
April 2013 to July 2014

<u>Country</u>
Scotland

<u>Setting</u>
Primary care</td><td><u>Patient recruitment</u>
All women, other than those previously diagnosed with CIN2+, attending routine screening in primary care were invited to consent to the study

<u>Data collection</u>
Participants first provided a random void urine sample, then self-collected a vaginal sample prior to a routine cervical sample being collected by a clinician.

Participants with high grade abnormalities were referred for colposcopy and women with borderline changes or low-grade cytology were recalled for repeat cytology after 6 months. Women with 2 low-grade or three borderline smears were referred for colposcopy.

<u>Sample size and demographics</u>
5,318 women

Mean age: 41.3 years (17 to 76)

Median age: 46 (<20=7, >59=145)</td><td><u>Self-testing</u>
Urine collected in universal containers, 6 ml was mixed with 3 ml of Roche PCR media.

Self-collected vaginal samples were obtained using cobas PCR female swab sample packets, women were advised to follow printed instructions, swabs were immediately immersed in tubes containing Roche PCR media.

Samples were tested with the cobas 4800 DNA HPV test using the standard procedure.

<u>Clinician-testing</u>
Cervical LBC samples were clinician collected using a Rovers Cervex-Brush and suspended in 20 ml of ThinPrep solution. 3 ml of this sample was aliquoted for HPV testing.

Samples were tested with the cobas 4800 DNA HPV test using the standard procedure.

negative cervical cytology results did not have underlying cervical precancer.

<u>Reference standard</u></td><td>Screening test accuracy

Comparison of self-testing (vaginal and urine samples) and clinician-collected cervical samples
<table>
<tr><th></th><th>Self-collected vaginal samples n=5208</th><th>Self-collected urine samples n=5003</th><th>Clinician collected cervical samples n=5299</th></tr>
<tr><td colspan="4">CIN2 or worse</td></tr>
<tr><td>Sensitivity (95% CI)</td><td>94.6% (90.7 to 98.5)</td><td>63.1% (54.6 to 71.7)</td><td>97.7% (95.0 to 100)</td></tr>
<tr><td>Specificity (95% CI)</td><td>85.4% (84.4 to 86.3)</td><td>59.8% (89.0 to 90.7)</td><td>87.3% (86.4 to 88.2)</td></tr>
<tr><td>LR+</td><td>6.48</td><td>1.57</td><td>7.69</td></tr>
<tr><td>LR-</td><td>0.054</td><td>0.62</td><td>0.026</td></tr>
<tr><td>Relative sensitivity (95% CI)</td><td>0.97 (0.94 to 1.00), p 0.1250</td><td>0.67 (0.59 to 0.76), p<0.0001</td><td>Reference</td></tr>
<tr><td>Relative specificity (95% CI)</td><td>0.98 (0.97 to 0.99), p<0.0001</td><td>1.05 (1.04 to 1.06), p<0.0001</td><td>Reference</td></tr>
<tr><td colspan="4">CIN3 or worse</td></tr>
<tr><td>Sensitivity</td><td>95.8% (91.1</td><td>50.7% (39.1</td><td>98.6% (95.9</td></tr>
</table>
</td></tr>
</table>

Study	Study design	Population characteristics	Screening methods	Epidemiology
		Of women aged ≤23 years, 66% (354/533) had been vaccinated with at least 2 doses of bivalent HPV vaccine	The reference standard for both tests was colposcopy. In the absence of a reference standard result it was assumed that women with no history of CIN2+ and 2 previous consecutively negative cervical cytology results did not have underlying cervical precancer.	

	to 100)	to 62.3)	to 100)
Specificity	84.8% (83.8 to 85.8)	89.7% (88.8 to 90.5)	86.4% (85.5 to 87.3)
LR+	6.30	4.92	7.25
LR-	0.045	0.49	0.016
Relative sensitivity (95% CI)	0.97 (0.93 to 1.01), p 0.5000	0.53 (0.42 to 0.67), p<0.0001	Reference
Relative specificity (95% CI)	0.98 (0.97 to 0.99), p<0.0001	1.03 (1.02 to 1.04) , p<0.0001	Reference

Table 16 Quality assessment of studies relevant to question 4

	Stanczuk 2016[36]
PARTICIPANT SELECTION	
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low
Is there concern that the included patients do not match the review question?	No
INDEX TEST	
Were the index test results interpreted without knowledge of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Unclear
Could the conduct or interpretation of the index test have introduced bias?	Low
Is there concern that the index test, its conduct, or interpretation differ from the review question?	No
REFERENCE STANDARD	
Is the reference standard likely to correctly classify the test condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low
Is there concern that the target condition as defined by the reference standard does not match the review question?	No
PARTICIPANT FLOW	
Was there an appropriate interval between the index test(s) and the reference standard?	Unclear
Did all patients receive a reference standard?	No

	Stanczuk 2016[36]
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High

Uptake of self-sampling

Details of studies relevant to question 5 are presented in Table 17, and full quality assessments are presented in Table 18.

Table 17 Studies relevant to question 5

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td>Verdoodt 2015[38] (SLR)</td>
<td><u>Design</u>
SLR and Meta-analysis

<u>Objective</u>
To evaluate if offering a kit for self-sampling (at home) could increase screening attendance, compared to sending reminder letters for a Pap smear or HPV test on a sample collected by a clinician (at the clinic)

<u>Dates</u>
Up to 12th February 2015</td>
<td><u>Eligible studies</u>
Population:
• Study population involved irregularly* or never-screened women, or women who did not respond to ≥1 invitation for conventional cervical cancer screening (collectively ‘under-screened women)
• A minimum of 1000 women were included in the study
*Women were considered to be irregularly screened if their last screening exceeded the locally defined screening interval

Intervention:
• Women in the interventional group (self-sampling arm) were invited to collect a self-sample for hrHPV testing
• Women in the control group were invited to undergo conventional cytology screening and/or hrHPV testing on a sample taken by a clinician
Outcome:
• The participation in the self-sampling arm was documented

<u>Data analysis</u>
Per protocol and ITT analyses performed, the latter includes data on women who were invited to perform a self-sample but instead opted to have a Pap smear test</td>
<td><u>Self-testing</u>
Kits were distributed by the following methods:
• Kit mailed directly to the home address of all women (Mail-to-all)
• Women were sent an invitation to order a self-sampling kit (Opt-in)
• Women were approached at their home and offered a self-sampling kit (Door-to-door)

<u>Clinician-testing</u>
In 12 studies, women in the control arm were invited for cytology.

In 2 studies there were arms for cytology and HPV testing.

In 2 studies only hrHPV testing was performed in the control arm.</td>
<td><u>Screening test uptake</u>
Comparison of self-testing and clinician-testing

*Low resource setting, women with limited access to health care as
<table>
<tr><th></th><th></th><th colspan="2">Participation</th></tr>
<tr><th></th><th>Studies, n</th><th>Self-sampling, % (95% CI)</th><th>Clinician-sampling, % (95% CI)</th></tr>
<tr><td colspan="4">Per-protocol</td></tr>
<tr><td>Mail-to-all</td><td>13</td><td>20.7
(16.9 to 24.8)
[range: 6.4% to 34.0%]</td><td>10.3
(6.2 to 15.2)</td></tr>
<tr><td>Opt-in</td><td>3</td><td>9.7
(6.5 to 13.5)</td><td>12.2
(10.9 to 13.6)</td></tr>
<tr><td>Door-to-door*</td><td>2</td><td>91.3
(65.8 to 100)</td><td>54.1
(0.9 to 100)</td></tr>
<tr><td colspan="4">ITT</td></tr>
<tr><td>Mail-to-all</td><td>13</td><td>23.6
(20.2 to 27.3)
[range: 10.2% to 39.0%]</td><td>10.3
(6.2 to 15.2)</td></tr>
<tr><td>Opt-in</td><td>3</td><td>14.0
(8.0 to 21.4)</td><td>12.2
(10.9 to 13.6)</td></tr>
<tr><td>Door-to-door*</td><td>2</td><td>92.4
(71.3 to 100)</td><td>54.1
(0.9 to 100)</td></tr>
</table>
opposed to having previously resistant to screening

Relative participation and participation difference in self-sampling vs</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td></td>
<td></td>
<td>
taken by a clinician.

Pooled proportions were calculated be a random effects model using metaprop, a statistical procedure for meta-analysis of binomial data. Relative rates and absolute differences were assessed by applying random effects models using metan.

<u>SLR results</u>
163,431 women in 16 studies included in the meta-analysis (2 studies in low resource settings)

134,262 women in 14 studies in high-resource settings

<u>Baseline characteristics</u>
'Under-screened':

Women who did not respond to 1 invitation for regular screening or an invitation and a reminder were eligible: 10 studies

Women who persistently did not respond to invitations for regular screening (more than 2 rounds of screening) were eligible: 4 studies

Women with limited access to health services in low-resource settings: 2 studies

<u>Quality assessment</u>
Study quality evaluated moderate to high, based on the Cochrane tool for bias

<table>
<tr><th rowspan="2">Risk of bias</th><th colspan="2">Selection</th><th>Attrition</th><th colspan="2">Reporting</th></tr>
<tr><th>S1</th><th>S2</th><th>A</th><th>R1</th><th>R2</th></tr>
<tr><td>Bais 2007</td><td>L</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Gok 2010</td><td>L</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Gorgi Rossi</td><td>L</td><td>L</td><td>L</td><td>L</td><td>L</td></tr>
</table>
</td>
<td></td>
<td>
control arm
^Two studies had 2 control arms (15 comparisons)

<table>
<tr><th></th><th>Studies, n</th><th>Relative participation, (95% CI)</th><th>Participation difference, % (95% CI)</th></tr>
<tr><td colspan="4">Per-protocol</td></tr>
<tr><td>Mail-to-all</td><td>13^</td><td>2.06
(1.44 to 2.96)</td><td>9.9
(5.8 to 13.9)</td></tr>
<tr><td>Opt-in</td><td>3†</td><td>0.72
(0.53 to 0.99)</td><td>-3.2 (-6.6 to 0.1)</td></tr>
<tr><td>Door-to-door</td><td>2</td><td>2.17
(0.33 to 14.13)</td><td>36.0
(-16.6 to 88.5)</td></tr>
<tr><td colspan="4">ITT</td></tr>
<tr><td>Mail-to-all</td><td>13^</td><td>2.40
(1.73 to 3.33)</td><td>12.6
(9.3 to 15.9)</td></tr>
<tr><td>Opt-in</td><td>3†</td><td>0.97
(0.65 to 1.46)</td><td>0.2
(-4.5 to 4.9)</td></tr>
<tr><td>Door-to-door</td><td>2</td><td>2.21
(0.32 to 15.48)</td><td>37.5
(-17.7 to 92.8)</td></tr>
</table>
†Two studies had 2 control arms (5 comparisons)

<u>Unsatisfactory tests</u>
Data on sample adequacy in the self-sampling are was reported in 12 studies.

The pooled proportion of unsatisfactory samples: 0.7% (95% CI 0.4 to 1.1%)
</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td></td>
<td></td>
<td>
<table>
<tr><td>2011</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Lazcano-Ponce 2011</td><td>M</td><td>L</td><td>L</td><td>M</td><td>M</td></tr>
<tr><td>Piana 2011</td><td>L</td><td>M</td><td>L</td><td>M</td><td>H</td></tr>
<tr><td>Szarewski 2011</td><td>M</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Virtanen 2011</td><td>L</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Wikstrom 2011</td><td>M</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Gok 2012</td><td>L</td><td>M</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Darlin 2013</td><td>M</td><td>M</td><td>L</td><td>M</td><td>M</td></tr>
<tr><td>Sancho-Garnier 2013</td><td>M</td><td>M</td><td>L</td><td>L</td><td>M</td></tr>
<tr><td>Broberg 2014</td><td>M</td><td>M</td><td>L</td><td>M</td><td>L</td></tr>
<tr><td>Cadman 2014</td><td>L</td><td>L</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Haguenoer 2014</td><td>L</td><td>L</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Arrossi 2015</td><td>M</td><td>L</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>Giorgi Rossi 2015</td><td>L</td><td>L</td><td>L</td><td>L</td><td>L</td></tr>
</table>
S1: Random sequence generation, S2: Allocation concealment, A: Incomplete outcome data, R1: Reporting of timelines, R2: Selective reporting

L: Low risk, M: Medium risk, H: high risk
</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bosgraaf 2015[39]

PROHTECT-3B</td>
<td><u>Design</u>
RCT

<u>Objective</u>
To assess the participation rate of 2 self-sampling methods and physician taken-smear, to compare performance of the sampling methods for detection of</td>
<td><u>Patient recruitment</u>
Women who did not respond to an invitation for a cervical smear in 2008, and living in the regions of North Holland Flevoland, Utrecht and Gelderland were invited to participate. All eligible women could opt-out.

Exclusion criteria:

• Previous hysterectomy
• Previous abnormal cytological test result within the last 2 years
• Current pregnancy

<u>Data collection</u></td>
<td><u>Self-testing</u>
Women were randomised (1:1) to receive either a brush device (Evalyn Brush) or a lavage device (second generation Delphi Screener) which was provided in a self-sampling kit with an explanatory letter, an informed consent form, user instructions, a questionnaire and a return envelope.</td>
<td><u>Screening test uptake</u>
Comparison of brush device and lavage device
<table>
<tr><th></th><th>Brush device, n=15,077</th><th>Lavage device, n=15,053</th><th>Eligible population, n=33,279</th></tr>
<tr><td>Participation rate, n (%, 95% CI]</td><td>5,218 (34.6, 33.9 to 35.4)</td><td>4,809 (31.9, 31.2 to 32.7)</td><td>10,027 (30.1)</td></tr>
<tr><td>Absolute difference, % (95% CI)</td><td colspan="2">2.7 (1.8 to 4.2)</td><td></td></tr>
</table>
Also reports participation rate in age ranges</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr><td></td><td>hrHPV and CIN2+ or CIN3+ and to investigate acceptability and user friendliness of the devices

<u>Dates</u>
Oct 2011 to Feb 2012

<u>Country</u>
Netherlands

<u>Setting</u>
Home testing - National Screening Programme</td><td>Women were sent self-sampling kits which they returned by post, all women who submitted self-samples between October 2011 and December 2012 were counted as self-sampling responders

<u>Sample size and demographics</u>
35,477 women were invited to take part, of these 33,279 were eligible, 5,347 (15.1%) opted out (3,149 of whom were eligible)

30,130 women randomised</td><td></td><td><u>Inadequate samples</u>
23/5,218 (0.4%) [95% CI 0.25 to 0.61] brush samples were inadequate for evaluation

24/4809 (0.5%) [95% CI 0.32 to 0.74] lavage samples were inadequate for evaluation</td></tr>
<tr><td>Ducancelle 2015[40]

CapU study</td><td><u>Design</u>
Prospective cohort study

<u>Objective</u>
To evaluate the participation rate of urinary HPV testing

<u>Dates</u>
July 2010 to Jan 2013

<u>Country</u>
France

<u>Setting</u>
Home testing</td><td><u>Patient recruitment</u>
Women aged 40 to 65 who had not responded to previous invitations and reminders for pap smears

<u>Exclusion criteria</u>
Previous hysterectomy

<u>Data collection</u>
Women accepting to participate returned a sample by mail to the Angers University Hospital Virology Laboratory, response rate determined on reception of informed consent forms and urine samples

<u>Sample size and demographics</u>
5,000 women aged 40 to 65

3,000 in a 40 to 54 years age group and 2,000 in a 55 to 65 years age group</td><td><u>Self-testing</u>
Women received an invitation letter with a urinary HPV DNA testing information note, a letter of consent a sterile container, a procedure protocol, a survey on the motives for refusal of the smear, a bubble envelope and a prepaid return envelope</td><td><u>Screening test uptake</u>
13.7% overall participation rate
<table><tr><th></th><th>40 to 54 years of age, n=3000</th><th>55 to 65 years of age, n=2000</th><th>p-value</th></tr><tr><td>Uptake, n (%)</td><td>512 (17)</td><td>259 (12.9)</td><td></td></tr><tr><td>Participation rate, n (%) [95% CI]</td><td>479 (15.9) [14.61 to 17.26]</td><td>208 (10.7) [9.38 to 12.14]</td><td><0.001</td></tr></table>Uptake represents number who returned samples, participation rate is a measure of the samples which were eligible for inclusion, women were considered ineligible due to prior hysterectomy or for refusal to participate in study.

<u>Inadequate samples</u>
• 40 to 54 years of age: 2 invalid tests (from eligible population)
• 55 to 65 years of age: 1 invalid test (from eligible population)
In total 3/687 (0.4%) [95% CI 0.07 to 1.22] samples could not be</td></tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr><td></td><td></td><td></td><td></td><td>analysed</td></tr>
<tr><td>Duke 2015[41]</td><td><u>Design</u>
Community based cohort study, case control

<u>Objective</u>
To determine whether offering self-collected HPV testing screening increased cervical cancer screening rates in rural communities

<u>Dates</u>
2010 to 2011

<u>Country</u>
Canada

<u>Setting</u>
Rural communities in the Canadian province of Newfoundland and Labrador</td><td><u>Patient recruitment</u>
All eligible women living in rural communities in the Canadian province of Newfoundland and Labrador

<u>Exclusion criteria</u>
Pregnancy

<u>Data collection</u>
In community A women were considered responders if they returned a self-collection kit, in community B women were considered responders if they presented for Pap smear and agreed to be part of the study; providing information about themselves and their screening history.

<u>Sample size and demographics</u>
1,760 women in community A, 2,761 women in community B and 1,536 women in community C at the end of the study</td><td><u>Self-testing</u>
Women in community A were given the opportunity of being screened for HPV infection through vaginal self-screening. A self-collection kit containing a Dacron swab, collection tube, instructions with explanatory pictures, consent forms and a participant questionnaire were available at public locations or a research nurse was available to drop off kits at a women's home or work. A kit was also available at the end of an educational presentation on cervical cancer screening.

<u>Clinician-testing</u>
In community B a campaign was used to raise awareness on the importance of regular screening through pap smear.

Women in community C received no intervention</td><td><u>Screening test uptake</u>
Return rate of self-collection kits

• 168/837 (20.1)
Comparison of response to self-testing and Pap smear invitation in eligible women

<table><tr><th></th><th>Community A: self-testing, n=1760</th><th>Community B: Pap smear, n=2761</th></tr><tr><td>Response rate, n (%) [95% CI]</td><td>168 (9.5) [7.80 to 11.43]</td><td>374 (13.5) [12.63 to 15.25]</td></tr><tr><td>Proportion of under- or unscreened participants, n (%)</td><td>26 (15.5)</td><td>52 (13.9)</td></tr></table>
Change in cervical cancer screening rates, 2008/2009 compared to 2010/2011

<table><tr><th></th><th>Community A</th><th>Community B</th><th>Community C</th></tr><tr><td>Cervical cancer screening rate 2008/2009, n/N[a] (%)</td><td>1,020/1,928 (52.9)</td><td>1,484/2,833 (52.4)</td><td>1,098/1,524 (72.0)</td></tr><tr><td>Cervical cancer screening rate 2010/2011, n/N[a] (%)</td><td>1,187/1,760 (67.4)</td><td>1,529/2,761 (55.3)</td><td>1,236/1,536 (80.5)</td></tr><tr><td>Change in rate, %</td><td>+15.2</td><td>+2.9</td><td>+8.5</td></tr><tr><td>p-value for change</td><td><0.001</td><td>0.07</td><td><0.01</td></tr><tr><td>p-value for difference in change</td><td>Reference</td><td><0.001</td><td>0.193</td></tr></table>
[a]Denominator (eligible population of women age 30 to 69) for the 2008/2009 period based on 2006 census data and for the 2010/2011 period is based on the 2011 census data

Cervical cancer screening rates for Women in Community A for 2010/2011 was determined as the number of women who had a pap smear and the number who did self-collection but did not have a pap smear</td></tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td>Enerly 2016[42]

SESAM study</td>
<td><u>Design</u>
RCT

<u>Objective</u>
To demonstrate the effects of self-sampling among women who do not attend the NCCSP, in particular:
• Impact of the self-sampling on screening attendance and coverage
• The performance of 2 different self-sampling devices for hrHPV testing
• Women's experience of the 2 self-sampling devices used

<u>Dates</u>
April/May 2013

<u>Country</u>
Norway

<u>Setting</u>
Home testing in National Screening Programme</td>
<td><u>Patient recruitment</u>
Non-attenders to the Norwegian Cervical Cancer Screening Programme (NCCSP), defined as a woman aged 26-69 years without any cytology, HPV or histology result recorded in the NCCSP registries within 12 months of the first reminder, identified in Oslo in April/May 2013

<u>Data collection</u>
Screening attendance was defined as either returning a self-sampling device and/or having a cervical smear taken by a clinician between April 2013 and the end of 2013

<u>Sample size and demographics</u>
3,393 women

800 assigned to the 'intervention group' (300 each from the age groups 26 to 34 and 35 to 49 years and 200 from the age group 50 to 69 years), 729 women were successfully contacted and consented to their inclusion in the trial

2,593 served as the control group</td>
<td><u>Self-testing</u>
Patients selected for the intervention group were sent an information letter inviting them to participate in the study. Those participating the study were randomized and sent 1 of 2 self-sampling devices along with user instructions, an informed consent form, a pre-paid return envelope and a questionnaire.

The self-sampling devices used in the study:
• Lavage based sampler: Delphi Screener™
• Dry brush sampler: Evalyn Brush

<u>Clinician-testing</u>
Followed according to the established procedures of the NCCSP; if no cytology result is recorded within 12 months of the initial reminder letter women were sent a second reminder, each woman is responsible for scheduling her own screening appointment</td>
<td><u>Screening test uptake</u>
Comparison of intervention and control group
<table>
<tr><th></th><th colspan="3">Intervention group</th><th>Control group</th></tr>
<tr><th></th><th>Delphi Screener n=400</th><th>Evalyn Brush n=400</th><th>Total n=800</th><th>Total n=2,593</th></tr>
<tr><td>Self-testing participants, n (%) [95% CI]</td><td>81 (20.3) [16.47 to 24.58]</td><td>88 (22.0) [18.04 to 26.38]</td><td>169 (21.1)</td><td></td></tr>
<tr><td>Cytology participants, n (%)</td><td>53 (13.3)</td><td>45 (11.3)</td><td>98 (12.3)</td><td></td></tr>
<tr><td>Total participants, n (%) [95% CI]</td><td>134 (33.5) [28.89 to 38.36]</td><td>133 (33.3) [28.70 to 38.15]</td><td>267 (33.4)</td><td>601 (23.2) [21.59 to 24.87]</td></tr>
</table>
ITT population

Attendance rates in the intervention and control groups are also reported by age (26 to 34, 35 to 49 and 50 to 69 years)

Total participation intervention/control arm relative risk: 1.44 (95% CI 1.28 to 1.62)

<u>Useable samples</u>
All 169 devices returned contained sufficient biological material for HPV testing</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td>Kitchener 2016[47]

STRATEGIC</td>
<td><u>Design</u>
RCT

<u>Objective</u>
To evaluate the clinical effectiveness of a range of interventions in:
• All women receiving their first invitation for cervical screening
• Those who had not attended by 6 months
To evaluate the cost effectiveness of these interventions and to study preferences for cervical screening among non-attenders.

<u>Dates</u>
April 2012 to June 2014 in north-west England

Oct 2012 to Dec 2014 in north-east Scotland

<u>Country</u>
UK (England and Scotland)

<u>Setting</u>
Home testing and</td>
<td><u>Patient recruitment</u>
Non-attenders to screening in Phase 1 of the trial during which women were sent their first routine invitation to attend cervical screening. Non-attenders were women who had no record of cytology test 6 months after their test date.

Women in Greater Manchester aged 24.75 years, women in Grampian aged 20 years.

<u>Data collection</u>
Data on uptake were obtained from the screening agency (Lancashire and South Cumbria Agency) in Greater Manchester and from the research team in Grampian, primary time point for uptake was 12 months following standard invitations

<u>Sample size and demographics</u>
10,126 women randomised (from 258 practices)
• 1,141 women received an unrequested self-sample kit
• 1,290 women received a letter offering a self-sample kit
• 3,782 women from 97 practices served as controls.
Patients were also randomised to be offered a nurse navigator (n=1,007), a timed appointment (n=1,629) or the option of a nurse navigator or self-sample kit (n=1,277)

<u>Baseline characteristics</u>
<table><tr><th>Characteristic</th><th>Women in Grampian, n=2,608</th></tr><tr><td colspan="2">Vaccination status</td></tr><tr><td>None</td><td>708</td></tr><tr><td>Incomplete</td><td>149</td></tr><tr><td>Full</td><td>1,724</td></tr><tr><td>Missing</td><td>27</td></tr></table></td>
<td><u>Self-testing</u>
There were two HPV self-sampling interventions:
• A letter offering the opportunity to request a self-sample kit
• An unrequested self-sample kit sent directly to the home
The self-sample kit comprised either a Delphi lavage device or The Rovers® Evalyn-Brush, an information sheet, a consent form and packaging to return the sample

<u>Clinician-testing</u>
Patients in the control arm were sent their first routine invitation for screening and received no further intervention</td>
<td><u>Screening test uptake</u>
Comparison of self-sampling kit sent, self-sampling kit offered and control groups (clinician sampling)
<table><tr><th>Intervention</th><th>Attendance, n (%) [95% CI]</th><th>OR[a] (95% CI)</th></tr><tr><td colspan="3">12 month follow up</td></tr><tr><td>Control, n=3,782</td><td>613 (16.2) [15.04 to 17.41]</td><td>Reference</td></tr><tr><td>Self-sampling sent, n=1,141</td><td>243 (21.3) [18.96 to 23.79]</td><td>1.512 (1.197 to 1.910), p=0.001</td></tr><tr><td>Self-sampling offered, n=1290</td><td>209 (16.2) [14.23 to 18.33]</td><td>1.074 (0.871 to 1.325), p=0.505</td></tr><tr><td colspan="3">18 month follow up</td></tr><tr><td>Control, n=3,782</td><td>27.1 (1026)</td><td>Reference</td></tr><tr><td>Self-sampling sent, n=1,141</td><td>30.0 (342)</td><td>1.286 (1.056 to 1.567), p=0.012</td></tr><tr><td>Self-sampling offered, n=1,290</td><td>25.8 (333)</td><td>1.056 (0.884 to 1.262), p=0.548</td></tr></table>[a] Adjusted OR associated with the change in odds of attendance occurring with intervention compared with control, adjusted for practice attendance rate and Primary Care Trust region

Attendance based on location, Greater Manchester or Grampian, is also reported

Type of screening undergone by participants
<table><tr><th></th><th colspan="4">Type of screen</th></tr><tr><th></th><th colspan="2">Single</th><th colspan="2">Both</th></tr><tr><th></th><th>HPV only</th><th>Cytology only</th><th>HPV first</th><th>Cytology first</th></tr><tr><td colspan="5">12 month follow up</td></tr><tr><td>Control, n=613</td><td>1</td><td>612</td><td>-</td><td>-</td></tr><tr><td>Self-sampling sent, n=243</td><td>52</td><td>158</td><td>32</td><td>1</td></tr><tr><td>Self-</td><td>12</td><td>190</td><td>7</td><td>-</td></tr></table></td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr><td></td><td>primary care</td><td></td><td></td><td>
<table>
<tr><td>sampling offered, n=209</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="5">18 month follow up</td></tr>
<tr><td>Control, n=1,026</td><td>1</td><td>1025</td><td>-</td><td>-</td></tr>
<tr><td>Self-sampling sent, n=342</td><td>59</td><td>248</td><td>34</td><td>1</td></tr>
<tr><td>Self-sampling offered, n=333</td><td>12</td><td>314</td><td>7</td><td>-</td></tr>
</table>
ITT population</td></tr>
<tr><td>Racey 2016[48]</td><td><u>Design</u>
RCT

<u>Objective</u>
To determine if cervical cancer screening uptake would increase among under-screened women living in rural Ontario, Canada, if at home self-collected sampling for HPV testing was offered as a primary cervical cancer screening modality, compared to invited Pap testing to routine opportunistic screening</td><td><u>Patient recruitment</u>
Women (aged 30 to 70 years) were identified as being under-screened/overdue for screening through their electronic medical record system, this was defined as not having had a pap test recorded in the preceding 30 months

<u>Exclusion criteria:</u>
• Residing in a long-term care facility
• Medical history of hysterectomy
• Any other medical contraindication
• Invalid mailing address
• Inactivated medical chart
Final eligibility was determined post-randomisation

<u>Data collection</u>
All women who participated in the self-collected HPV test had their results recorded in their medical chart.

Pap test completion was recorded from the medical charts at the end of the study period for eligible women in the study</td><td><u>Self-testing</u>
Women were sent a study information letter informing them about the study and giving them the option to opt-out 2 weeks before the self-collection kit was sent.

The self-collection kit contained a vaginal swab, collection tube, annotated pictorial instructions, a questionnaire, an information sheet on cervical cancer and HPV and a return envelope.

A reminder phone call was placed to non-responders 1 month after self-collection kits were sent

<u>Clinician-testing</u>
Women in the Pap testing arm were sent an invitation letter that asked for them to call their doctor and book an</td><td><u>Screening test uptake</u>
Comparison of self-testing, pap testing and opportunistic screening
<table>
<tr><th colspan="2"></th><th>Screened, n (%) [95% CI)</th></tr>
<tr><td rowspan="3">Self-collection arm, n=335</td><td>Self-collection sample</td><td>70 (21) [16.76 to 25.76]</td></tr>
<tr><td>Pap testing</td><td>37 (11)</td></tr>
<tr><td>Total</td><td>107 (32) [27.03 to 37.29]</td></tr>
<tr><td colspan="2">Pap invitation arm, n=331</td><td>51 (15.4) [11.33 to 19.31]</td></tr>
<tr><td colspan="2">Standard of care, n=152</td><td>13 (8.6) [8.10 to 19.41]</td></tr>
</table>
• Women in self-collection arm were 3.7 (95% CI 2.2 to 6.4) times more likely to undergo screening compared with the standard of care arm
• Women in Pap test arm were 1.8 (95% CI 1.0 to 3.2) times more likely to screen compared to women in the standard of care arm
• Women in self-collection arm were 2.1 (95% CI 1.5 to 2.8) times more likely to undergo screening compared with women in the Pap test arm (p=0.097)</td></tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td></td>
<td><u>Dates</u>
October 2012 to July 2013

<u>Country</u>
Canada

<u>Setting</u>
Rural community where low rates of cervical cancer screening have been observed, in partnership with primary care</td>
<td>A modified ITT analysis was used for all post-randomisation eligible women to calculate the RR for each arm.

<u>Sample size and demographics</u>
964 women identified as under-screened and randomised:

• 400 to self-collection
• 400 to Pap invitation
• 164 to standard of care opportunistic screening- women seeking cervical cancer screening through their own initiative, with or without prompting from a healthcare provider
After adjusting for eligibility:

• 818 eligible women
• 335 received a self-collected HPV testing kit
• 331 received a reminder letter
• 152 received standard of care opportunistic screening
No women contacted the clinic to opt-out

<u>Baseline characteristics</u>
<table>
<tr><th>Characteristic</th><th>Self-collection arm</th><th>Pap test arm</th></tr>
<tr><td>Mean age, years (95% CI)</td><td>53.6 (51.2 to 56.0)</td><td>50.5 (46.0 to 55.0)</td></tr>
<tr><td>Age, years, n (%)</td><td>n=76</td><td>n=24</td></tr>
<tr><td>30 to 39</td><td>7 (9.2)</td><td>3 (13.0)</td></tr>
<tr><td>40 to 49</td><td>22 (29.0)</td><td>8 (34.8)</td></tr>
<tr><td>50 to 59</td><td>18 (23.7)</td><td>9 (39.1)</td></tr>
<tr><td>60+</td><td>29 (38.2)</td><td>4 (16.7)</td></tr>
<tr><td>Screening history, n (%)</td><td>n=76</td><td>n=23</td></tr>
<tr><td>Prior Pap test, yes</td><td>75 (98.7)</td><td>23 (100)</td></tr>
<tr><td>3 years or more since last Pap test</td><td>47 (62.7)</td><td>14 (60.9)</td></tr>
</table>
</td>
<td>appointment, in addition to an information sheet HPV and cervical cancer screening. Women who did not respond within 1 month were called by the clinic to follow-up and book an appointment if possible, a change in the protocol during the trial led to only 20% of the women in the Pap invitation arm receiving a follow up call due to a shortage in resources.

Women in the opportunistic screening arm were not contacted during the study period.</td>
<td>Underpowered exploratory sub analysis (per protocol):

• Uptake of self-collected sampling for HPV testing vs the standard of care arm: RR= 2.4 (95% CI 1.4 to 4.3), significantly higher
• Uptake of self-collected sampling for HPV testing vs Pap test arm: RR= 1.4 (95% CI 0.98 to 1.9), no significant difference

<u>Adequate samples</u>
1/70 (1.4%) [95% CI 0.03 to 7.66] samples were not β-globin positive, which demonstrates a high DNA sample quality</td>
</tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><td><2 years since last Pap test</td><td>24 (32.0)</td><td>5 (21.7)</td></tr>
<tr><td>Do not remember</td><td>4 (5.3)</td><td>4 (17.4)</td></tr>
</table>
</td><td></td><td></td></tr>
<tr><td>Sultana 2016[43]

iPap</td><td><u>Design</u>
RCT

<u>Objective</u>
To determine if HPV self-sampling could increase participation in the Australian cervical cancer screening program

<u>Dates</u>
March to July 2014

<u>Country</u>
Australia

<u>Setting</u>
Home testing and testing in primary care, data from Pap test registers</td><td><u>Patient recruitment</u>
Women who were residents of Victoria and were identified through the Victorian Cervical Cytology Register as never-screened (women on the electoral role but for whom no match was found on the registry) or under-screened (not screened in the past 5 years)

<u>Eligibility criteria:</u>
• Aged 30 to 69 years
• Not pregnant
• Not had a hysterectomy

<u>Data collection</u>
Primary outcome was participation in screening at three and 6 months after initial letters were mailed, indicated by returning a self-sampling swab or having a Pap test, women who had Pap tests after randomisation were identified by performing a semi-automated match of the trial database with Registry records of Pap tests conducted in 2014

<u>Sample size and demographics</u>
8,160 women

7,140 in the self-sampling arm and 1,020 in the Pap test arm

<u>Baseline characteristics</u>
<table>
<tr><th></th><th colspan="2">Apparently never screened</th><th colspan="2">Apparently under-screened</th></tr>
<tr><th></th><th>S, n= 7,140</th><th>P, n= 1,020</th><th>S, n= 7,140</th><th>P, n= 1,020</th></tr>
<tr><td colspan="5">Age, years n (%)</td></tr>
<tr><td>30 to 39</td><td>1,950 (27.3)</td><td>276 (27.1)</td><td>2,334 (32.7)</td><td>323 (31.7)</td></tr>
</table>
</td><td><u>Self-testing</u>
Women randomised to self-sampling arm were sent a pre-invitation letter, informing them that they would be receiving a kit and giving the opportunity to withdraw from the trial. Participants were then sent a package containing an information brochure, a nylon-tipped flocked swab enclosed in a dry plastic tube within a re-sealable plastic bag, an instruction sheet, a personal information form and a postage paid envelope

<u>Clinician-testing</u>
Women received a single invitation letter (never screened population) or a standard reminder letter (under-screened population) to have a pap test, this included a pap test brochure, a personal information form and a postage paid envelope</td><td><u>Screening test uptake</u>
Comparison of response within 6 months to self-testing and clinician-testing invitations in never-screened and under-screened population
<table>
<tr><th></th><th colspan="3">Self-sampling arm, n=7,140</th><th>Pap test arm, n=1,020</th><th rowspan="2">Absolute difference (95% CI)</th></tr>
<tr><th></th><th>Self-sampling</th><th>Pap test</th><th>Total</th><th>Pap test</th></tr>
<tr><td>Never screened, n (%) [95% CI]</td><td>1,131 (15.8) [14.96 to 16.67]</td><td>321 (4.5)</td><td>1,452 (20.3) [19.37 to 21.25]</td><td>61 (6) [4.62 to 7.64]</td><td>14.4% (12.6 to 16.1, p<0.001)</td></tr>
<tr><td>Under-screened, n (%) [95% CI]</td><td>518 (7.3) [6.71 to 7.93]</td><td>300 (4.2)</td><td>818 (11.5) [10.77 to 12.26]</td><td>65 (6.4) [4.98 to 8.08]</td><td>5.1% (3.4 to 6.8, p<0.001)</td></tr>
</table>
A sensitivity analysis was performed for the never-screened population to account for women who were determined ineligible for the trial post-randomisation. This was either women who were found to have had a prior Pap test or women who were found to have had a prior hysterectomy. The results from these analyses determined participation rates of 14.2% for the self-sampling arm and 4.2% for the Pap test arm, with an absolute difference of 10%.

Difference in participation between arms stratified by age, socioeconomic status and time from last Pap test also reported.

<u>Unsatisfactory tests</u>
9 (0.6%) [95% CI 0.29 to 1.10] of the returned samples were found to be unsatisfactory</td></tr>
</table>

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><td>40 to 49</td><td>1,342 (18.8)</td><td>176 (17.3)</td><td>2,351 (32.9)</td><td>358 (35.1)</td></tr>
<tr><td>50 to 59</td><td>1,453 (20.4)</td><td>198 (19.4)</td><td>1,453 (20.4)</td><td>207 (20.3)</td></tr>
<tr><td>60 to 69</td><td>2,395 (33.5)</td><td>370 (36.3)</td><td>1,002 (14.0)</td><td>132 (12.9)</td></tr>
</table>
Baseline characteristics of socioeconomic status and area remoteness also reported
</td><td></td><td></td></tr>
<tr><td>Tamalet 2016[44]</td><td><u>Design</u>
Prospective cohort

<u>Objective</u>
To describe high risk-HPV types in 35 to 69-year-old women from low socioeconomic groups not attending regular cytological screening in Marseille, France

<u>Dates</u>
2011 to 2012

<u>Country</u>
France

<u>Setting</u>
Home testing</td><td><u>Patient recruitment</u>
Women aged 35 to 69 years living in the Northern districts of Marseille were identified in the National Insurance Registry as not having had a Pap smear for more than 2 years. Women were informed by mail the importance of regular screening and that they would receive a vaginal self-sampling kit at home in the next month

<u>Data collection</u>
Women who returned a self-sample were considered participants.

<u>Sample size and demographics</u>
27,000 women initially contacted, 22,702 were sent self-sampling HPV tests after elimination of women refusing tests or not living at the mailing address

<u>Baseline characteristics</u>
<table>
<tr><th>Characteristic</th><th>Eligible population</th></tr>
<tr><td colspan="2">Age, years</td></tr>
<tr><td>35 to 39, n</td><td>4,395</td></tr>
<tr><td>40 to 44, n</td><td>4,211</td></tr>
<tr><td>45 to 49, n</td><td>4,044</td></tr>
<tr><td>50 to 54, n</td><td>3,488</td></tr>
<tr><td>55 to 59, n</td><td>2,886</td></tr>
<tr><td>60 to 64, n</td><td>1,600</td></tr>
<tr><td>65 to 69, n</td><td>2,078</td></tr>
</table>
</td><td><u>Self-testing</u>
Women were sent a self-sampling HPV test with instructions and a response envelope. Vaginal cells and secretions were collected using flocked swabs (MAST diagnostics) and subsequently placed in Abbott transport medium. The swab was placed in a tube and then sent in the mail to a Virology laboratory participating in the study.</td><td><u>Screening test uptake</u>
4,245/22,702 (18.7%) [95% CI 18.19 to 19.21] women performed self-sampling

Participation is also reported by age groups

<u>Unsatisfactory tests</u>
9/4,245 (0.21%) [95% CI 0.10 to 0.40] samples were excluded due to low cellularity</td></tr>
</table>

Study	Study design	Population characteristics	Screening methods	Results
Verhoef 2014[45]	Design RCT Objective To investigate if direct DNA methylation-based molecular triage on self-sampled cervicovaginal specimens was non-inferior to cytology triage on additional physician-collected cervical samples in the detection of CIN2+ in women who did not attend cervical screening programmes Dates 1 Nov 2010 to 31 Dec 2011 Country Netherlands Setting Home testing, centres: VU University Medical Centre, Radboud University Nijmegen Medical Centre, the screening organisations Mid-West and East	Patient recruitment Women registered as non-attendees in 2007 in the databases of screening organisations Eligibility criteria: • Aged 33 to 63 years • Living in Noord-Holland Flevoland, Utrecht and Gelderland • No hysterectomy • No history of CIN2+ • No abnormal cytology in the preceding 2 years Data collection Women who returned samples and informed consent forms were considered responders Sample size and demographics 46,001 women invited, 38,913 sent self-sampling devices	Self-testing Non-attendees were sent a letter allowing them to opt out of the trial, and those who did not opt out subsequently received a self-sampling lavage device (Delphi screener), an explanation letter, an informed consent form, an instruction form, a collection tube, a seal bag and a free return envelope. Women were asked to return their self-sampled material, together with a signed consent form to the laboratory for hrHPV testing.	Screening test uptake 12,819/38,913 (*32.9%*) self-sampling devices returned 12,819/46,001 (*27.9%*) [95% CI *27.49 to 28.*31] women invited to take part in study

<table>
<tr><th>Study</th><th>Study design</th><th>Population characteristics</th><th>Screening methods</th><th>Results</th></tr>
<tr>
<td>Virtanen 2015[46]</td>
<td><u>Design</u>
Prospective cohort

<u>Objective</u>
To study the effect of reminder letters (1st reminder) and self-sampling tests (2nd reminder) as means to increase attendance within the routine cervical cancer screening programme

<u>Dates</u>
2011-2012

(11 municipalities took part both years, 11 only in 2011 and 9 only in 2012)

<u>Country</u>
Finland

<u>Setting</u>
Home testing</td>
<td><u>Patient recruitment</u>
Women were identified for screening from the Population Register based on their age and home municipality, and all with address information available are invited to screening by personal letters. Non-attendees received a second invitation (1st reminder) within the same year, however in 2012, women were not sent a reminder letter if they cancelled their given appointment. As a second reminder letter, a self-sampling test was sent out to non-attendees. Prior to mailing the device, the possibility was introduced in an invitation letter with an opt out option.

<u>Data collection</u>
Women were considered attenders by returning a self-taken sample of by coming to the clinic for a Pap smear

<u>Sample size and demographics</u>
31,053 women identified for screening, of whom 30,827 received an initial invitation to screening.

4,536 invited to obtain self-sampling kit, of whom 3,836 received the kit

<u>Characteristics of women invited to self-sampling</u>
<table>
<tr><th>Characteristic</th><th>Women invited to self-sampling, n=4536</th></tr>
<tr><td colspan="2">Age, years</td></tr>
<tr><td>30 to 34, n</td><td>994</td></tr>
<tr><td>35 to 39, n</td><td>753</td></tr>
<tr><td>40 to 44, n</td><td>528</td></tr>
<tr><td>45 to 49, n</td><td>585</td></tr>
<tr><td>50 to 54, n</td><td>535</td></tr>
<tr><td>55 to 59, n</td><td>562</td></tr>
<tr><td>60 to 64, n</td><td>579</td></tr>
</table>
Characteristics of mother tongue, municipality type, education level, marital status, geographical location are also reported</td>
<td><u>Self-testing</u>
The sample taking was done by the Delphi Scanner (lavage device). Samples were sent to the screening laboratory in a test-tube in the regular mail.</td>
<td><u>Screening test uptake</u>
939/4,536 (20.7%) [95% CI 19.53 to 21.91] women took part in screening after invitation to receive a self-sampling kit

920 (20.3%) [95% CI 19.14 to 21.50] returned a self-sampling kit

19 (0.4%) attended a Pap-smear

Increase in total participation rate to 82.2% (95% CI: 81.8 to 82.7) from 79.2% (95% CI 78.8 to 79.7) after 1st reminder

Screening attendance by age-group, mother tongue, municipality type, education level, marital status and geographical location also reported

<u>Unsatisfactory tests</u>
30/920 (3.2%) [95% CI 2.16 to 4.55] of the originally returned samples were not considered adequate.

(Only samples which produced a visible pellet after centrifugation at 1500 rpm were considered adequate)</td>
</tr>
</table>

Study	Study design	Population characteristics	Screening methods	Results
		Characteristics of self-sampling participants Time from previous Pap-smear, years / Self-sampling participants, n=939 <5, n (%) 533 (56.8) 5 to 9, n (%) 157 (16.7) ≥10 years, n (%) 72 (7.7) Never, n (%) 40 (4.3) No information, n (%) 137 (14.6)		

Table 18 Quality assessment of studies relevant to question 5

Question	Bosgraaf 2015[39]	Ducancelle 2015[40]	Duke 2015[41]	Enerly 2016[42]	Sultana 2016[43]	Tamalet 2016[44]	Verhoef 2014[45]	Virtanen 2015[46]	Kitchener 2016[47]	Racey 2016[48]
EXTERNAL VALIDITY										
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	No	Yes	Yes	No	Yes	No	No	No	Yes	Yes
Were the subjects asked to participate in the study representative of the population of interest for this review?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CONFOUNDING										
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients?	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the statistical tests used to assess the main outcomes appropriate?	Yes	N/A	Yes	Yes	Yes	N/A	N/A	N/A	Yes	Yes
Were the main outcome measures used accurate (valid and reliable)?	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes

Question	Bosgraaf 2015[39]	Ducancelle 2015[40]	Duke 2015[41]	Enerly 2016[42]	Sultana 2016[43]	Tamalet 2016[44]	Verhoef 2014[45]	Virtanen 2015[46]	Kitchener 2016[47]	Racey 2016[48]
Were the patients in different intervention groups (trials and cohort studies) recruited from the same population?	Yes	N/A	No	Yes	Yes	N/A	N/A	N/A	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Unclear	N/A	Unclear	Unclear	Yes	N/A	N/A	N/A	Unclear	Yes
Were study subjects in different intervention groups (trials and cohort studies) recruited over the same period of time?	Yes	N/A	Yes	Yes	Yes	N/A	N/A	N/A	Yes	Yes
Were study subjects randomised to intervention groups?	Yes	No	No	Inadequate	Yes	No	No	No	Yes	Yes
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	Yes	N/A	N/A	Unclear	Yes	N/A	N/A	N/A	Yes	Unclear
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Yes	N/A	No	Unclear	Yes	N/A	N/A	N/A	Unclear	Yes
POWER										
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	No	N/A	Unclear	Unclear	Unclear	N/A	N/A	N/A	Yes	Unclear

Appendix 5 – Explanation of screening test accuracy graphs

[To be completed after finalisation of graph design and development of wording for 'Examples and Explanations' document]

Lines represent values of sensitivity and specificity that give the same LR+ or LR-

Figure 11 Positive likelihood ratios

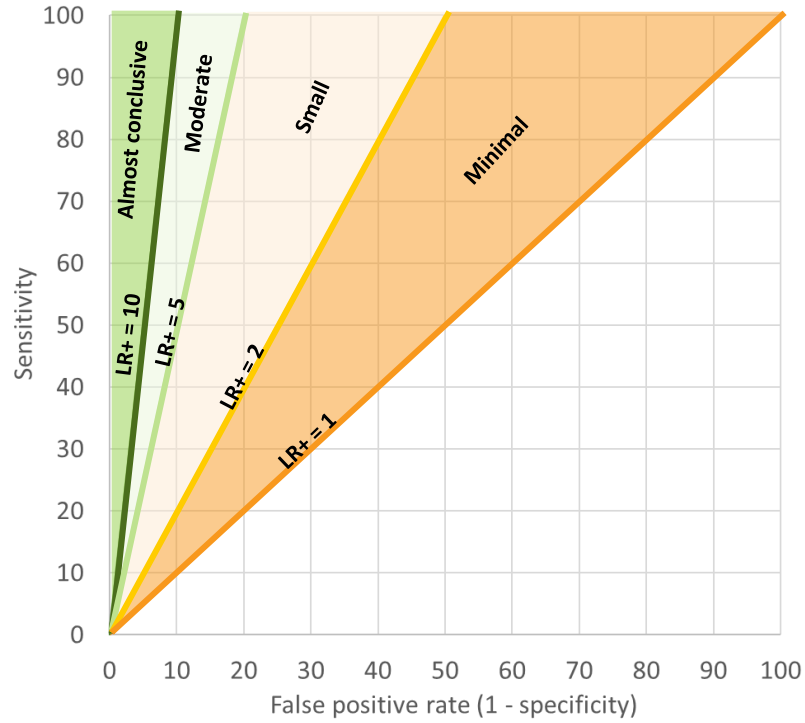
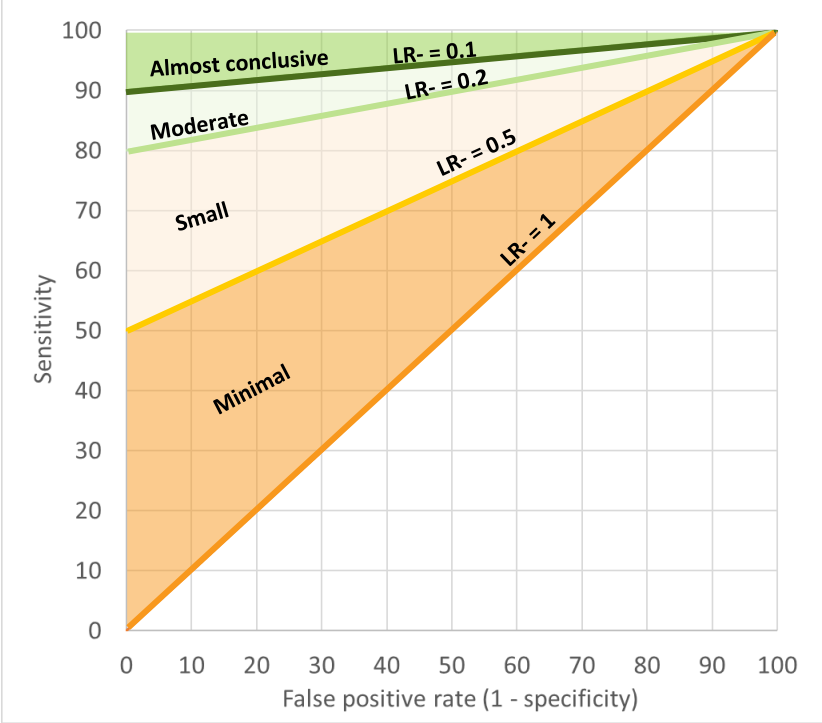


Figure 12 Negative likelihood ratios



5 References

1. Ferlay J, Soerjomataram I, Ervik M, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr. Last accessed 4th March 2017. 2013.
2. Cancer Research UK. Cervical cancer mortality statistics. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer/mortality. Last accessed 5th March 2017., 2014.
3. Nayar R, Wilbur DC. The Pap test and bethesda 2014. Acta cytologica 2015;59:121-132.
4. Cancer Research UK. Treatment if you have abnormal cervical cells. Available from: http://www.cancerresearchuk.org/about-cancer/cervical-cancer/treatment-for-abnormal-cervical-cells/treatment. Last accessed 5th March 2017., 2014.
5. Franco EL, Villa LL, Sobrinho JP, et al. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. Journal of Infectious Diseases 1999;180:1415-1423.
6. Cancer Research UK. Cervical cancer statistics. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer. Last accessed 5th March 2017., 2014.
7. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. The Journal of pathology 1999;189:12-19.
8. Chesson HW, Dunne EF, Hariri S, et al. The estimated lifetime probability of acquiring human papillomavirus in the United States. Sexually transmitted diseases 2014;41:660-664.
9. Burchell AN, Winer RL, de Sanjose S, et al. Chapter 6: Epidemiology and transmission dynamics of genital HPV infection. Vaccine 2006;24 Suppl 3:S3/52-61.
10. Anderson L, O'Rorke M, Jamison J, et al. Prevalence of human papillomavirus in women attending cervical screening in the UK and Ireland: new data from northern Ireland and a systematic review and meta-analysis. J Med Virol 2013;85:295-308.
11. Hibbitts S, Tristram A, Beer H, et al. UK population based study to predict impact of HPV vaccination. J Clin Virol 2014;59:109-14.
12. Ho GY, Bierman R, Beardsley L, et al. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med 1998;338:423-8.
13. Baseman JG, Koutsky LA. The epidemiology of human papillomavirus infections. Journal of clinical virology 2005;32:16-24.
14. Richardson H, Kelsall G, Tellier P, et al. The natural history of type-specific human papillomavirus infections in female university students. Cancer Epidemiology and Prevention Biomarkers 2003;12:485-490.
15. Public Health England. Cervical screening: programme overview, https://www.gov.uk/guidance/cervical-screening-programme-overview. Last accessed: 4th March 2017., 2015.
16. Kitchener HC. Report to the National Screening Committee. Available from: https://legacyscreening.phe.org.uk/policydb_download.php?doc=555. Last accessed 4th March 2017. 2015.
17. UK National Screening Committee. UK NSC HPV recommendation. Available from: https://legacyscreening.phe.org.uk/policydb_download.php?doc=594. Last accessed 4th March 2017. 2016.

18. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine 2012;30 Suppl 5:F88-99.
19. Health and Social Care Information Centre. Cervical Screening Programme, England. Statistics for 2014-15. Available from: http://content.digital.nhs.uk/catalogue/PUB18932/nhs-cervical-stat-eng-2014-15-rep.pdf. Last accessed 14th March 2017. 2015.
20. Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: Joint European cohort study. Bmj 2008;337:969-972.
21. Kitchener HC, Almonte M, Thomson C, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. The lancet oncology 2009;10:672-682.
22. Cook DA, Mei W, Smith LW, et al. Comparison of the Roche cobas® 4800 and Digene Hybrid Capture® 2 HPV tests for primary cervical cancer screening in the HPV FOCAL trial. BMC cancer 2015;15:968.
23. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. The lancet oncology 2006;7:547-555.
24. Anttila A, Kotaniemi-Talonen L, Leinonen M, et al. Rate of cervical cancer, severe intraepithelial neoplasia, and adenocarcinoma in situ in primary HPV DNA screening with cytology triage: randomised study within organised screening programme. BMJ (Clinical research ed.). Volume 340, 2010:c1804.
25. Malila N, Leinonen M, Kotaniemi-Talonen L, et al. The HPV test has similar sensitivity but more overdiagnosis than the Pap test - A randomised health services study on cervical cancer screening in Finland. International Journal of Cancer 2013;132:2141-2147.
26. Leinonen MK, Nieminen P, Lonnberg S, et al. Detection rates of precancerous and cancerous cervical lesions within one screening round of primary human papillomavirus DNA testing: Prospective randomised trial in Finland. BMJ (Online) 2012;345 (7886) (no pagination).
27. Veijalainen O, Kares S, Kujala P, et al. Human papillomavirus test with cytology triage in organized screening for cervical cancer. Acta Obstetricia et Gynecologica Scandinavica 2016;95:1220-1227.
28. Del Mistro A, Frayle H, Ferro A, et al. Cervical cancer screening by high risk HPV testing in routine practice: results at one year recall of high risk HPV-positive and cytology-negative women. Journal of medical screening 2014;21:30-37.
29. Zorzi M, Del Mistro A, Farruggio A, et al. Use of a high-risk human papillomavirus DNA test as the primary test in a cervical cancer screening programme: A population-based cohort study. BJOG: An International Journal of Obstetrics and Gynaecology 2013;120:1260-1268.
30. Zorzi M, Frayle H, Rizzi M, et al. A 3-year interval is too short for re-screening HPV negative women: a population-based cohort study. Bjog 2017.
31. Gustinucci D, Rossi PG, Cesarini E, et al. Use of Cytology, E6/E7 mRNA, and p16INK4a–Ki-67 to Define the Management of Human Papillomavirus (HPV)–Positive Women in Cervical Cancer Screening. American journal of clinical pathology 2016;145:35-45.
32. Gök M, Heideman DA, van Kemenade FJ, et al. HPV testing on self collected cervicovaginal lavage specimens as screening method for women who do not attend cervical screening: cohort study. Bmj 2010;340:c1040.

33. Naucler P, Ryd W, Törnberg S, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. Journal of the National Cancer Institute 2009;101:88-99.
34. Rijkaart D, Berkhof J, Van Kemenade F, et al. HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. British journal of cancer 2012;106:975-981.
35. Arbyn M, Verdoodt F, Snijders PJ, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. The lancet oncology 2014;15:172-183.
36. Stanczuk G, Baxter G, Currie H, et al. Clinical validation of hrHPV testing on vaginal and urine self-samples in primary cervical screening (cross-sectional results from the Papillomavirus Dumfries and Galloway-PaVDaG study). BMJ Open 2016;6:e010660.
37. Belinson JL, Du H, Yang B, et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. Int J Cancer 2012;130:1855-60.
38. Verdoodt F, Jentschke M, Hillemanns P, et al. Reaching women who do not participate in the regular cervical cancer screening programme by offering self-sampling kits: a systematic review and meta-analysis of randomised trials. European journal of cancer 2015;51:2375-2385.
39. Bosgraaf RP, Verhoef VMJ, Massuger LFAG, et al. Comparative performance of novel self-sampling methods in detecting high-risk human papillomavirus in 30,130 women not attending cervical screening. International Journal of Cancer 2015;136:646-655.
40. Ducancelle A, Reiser J, Pivert A, et al. Home-based urinary HPV DNA testing in women who do not attend cervical cancer screening clinics. Journal of Infection 2015;71:377-384.
41. Duke P, Godwin M, Ratnam S, et al. Effect of vaginal self-sampling on cervical cancer screening rates: A community-based study in Newfoundland. BMC Women's Health 2015;15 (1) (no pagination).
42. Enerly E, Bonde J, Schee K, et al. Self-Sampling for Human Papillomavirus Testing among Non-Attenders Increases Attendance to the Norwegian Cervical Cancer Screening Programme. PLoS ONE [Electronic Resource] 2016;11:e0151978.
43. Sultana F, English DR, Simpson JA, et al. Home-based HPV self-sampling improves participation by never-screened and under-screened women: Results from a large randomized trial (iPap) in Australia. International Journal of Cancer 2016;139:281-290.
44. Tamalet C, Halfon P, Retraite LL, et al. Genotyping and follow-up of HR-HPV types detected by self-sampling in women from low socioeconomic groups not participating in regular cervical cancer screening in France. Journal of Clinical Virology 2016;78:102-107.
45. Verhoef VMJ, Bosgraaf RP, Van Kemenade FJ, et al. Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): A randomised controlled non-inferiority trial. The Lancet Oncology 2014;15:315-322.
46. Virtanen A, Anttila A, Luostarinen T, et al. Improving cervical cancer screening attendance in Finland. International Journal of Cancer 2015;136:E677-E684.
47. Kitchener HC, Gittins M, Rivero-Arias O, et al. A cluster randomised trial of strategies to increase cervical screening uptake at first invitation (STRATEGIC). Health Technology Assessment (Winchester, England) 2016;20:1-138.
48. Racey CS, Gesink DC, Burchell AN, et al. Randomized Intervention of Self-Collected Sampling for Human Papillomavirus Testing in Under-Screened Rural Women: Uptake of Screening and Acceptability. Journal of Women's Health 2016;25:489-497.

49. Berkhof J, Coupe VM, Bogaards JA, et al. The health and economic effects of HPV DNA screening in The Netherlands. Int J Cancer 2010;127:2147-58.
50. Naber SK, Matthijsse SM, Rozemeijer K, et al. Cervical Cancer Screening in Partly HPV Vaccinated Cohorts - A Cost-Effectiveness Analysis. PLoS One 2016;11:e0145548.
51. Ronco G, Biggeri A, Confortini M, et al. [Health technology assessment report: HPV DNA based primary screening for cervical cancer precursors]. Epidemiol Prev 2012;36:e1-72.
52. Rozemeijer K, de Kok IM, Naber SK, et al. Offering Self-Sampling to Non-Attendees of Organized Primary HPV Screening: When Do Harms Outweigh the Benefits? Cancer Epidemiol Biomarkers Prev 2015;24:773-82.
53. Public Health England. UK NSC evidence review process. Guidance. Appendix F: Requirements for UK NSC evidence summaries. Available from: https://www.gov.uk/government/publications/uk-nsc-evidence-review-process/appendix-f-requirements-for-uk-nsc-evidence-summaries. Last accessed 4th March 2017. 2016.
54. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
55. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
56. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of Epidemiology and Community Health 1998;52:377-384.
57. Cook DA, Mei W, Smith LW, et al. Comparison of the Roche cobas 4800 and Digene Hybrid Capture 2 HPV tests for primary cervical cancer screening in the HPV FOCAL trial. BMC Cancer 2015;15 (1) (no pagination).
58. Gustinucci D, Rossi PG, Cesarini E, et al. Use of Cytology, E6/E7 mRNA, and p16INK4a-Ki-67 to Define the management of human papillomavirus (HPV)-positive women in cervical cancer screening. American Journal of Clinical Pathology 2016;145:35-45.
59. Veijalainen O, Kares S, Kujala P, et al. HPV test with cytology triage in organized screening for cervical cancer. Acta Obstetricia et Gynecologica Scandinavica 2016;3:3.
60. Zorzi M, Del Mistro A, Farruggio A, et al. Use of a high-risk human papillomavirus DNA test as the primary test in a cervical cancer screening programme: a population-based cohort study. BJOG: An International Journal of Obstetrics & Gynaecology 2013;120:1260-1268.