National Cervical Screening Program Renewal:

Evidence review

November 2013

MSAC application no 1276

Assessment report

ISBN (Online) *<number>*

ISSN (Online) 1443-7139

Electronic copies of the report can be obtained from the Medical Service Advisory Committee's internet site at http://www.msac.gov.au/

Enquiries about the content of the report should be directed to:

The Secretary
Medical Services Advisory Committee
Department of Health and Ageing
Mail Drop 106
GPO Box 9848
Canberra ACT 2601

The Medical Services Advisory Committee (MSAC) is an independent committee which has been established to provide advice to the Minister for Health and Ageing on the strength of evidence available on new and existing medical technologies and procedures in terms of their safety, effectiveness and cost-effectiveness. This advice will help to inform government decisions about which medical services should attract funding under Medicare.

MSAC's advice does not necessarily reflect the views of all individuals who participated in the MSAC evaluation.

Publication approval number: *<number>*

Template Version updated Nov 08

Contents

Table of Contents

List of Summaries

List of Tables

List of Figures

Introduction

In 2011, the former Australian Population Health Development Principal Committee of the Australian Health Ministers' Advisory Council endorsed the plan to renew the National Cervical Screening Program (NCSP). The Renewal of the NCSP commenced late 2011 and is planned for completion by mid-2014.

The Renewal aims to ensure the continuing success of the NCSP and that all Australian women, HPV-vaccinated and unvaccinated, have access to a cervical screening program based on current evidence and best practice. The Renewal process will ensure the renewed program remains safe and continues to improve health outcomes of Australian women.

The objectives of the Renewal are to:

1. Assess the evidence for screening tests and pathways, the screening interval, age range and commencement for both vaccinated and non-vaccinated women.
2. Determine a cost-effective screening pathway and program model.
3. Investigate options for improved national data collection systems and registry functions to enable policy, planning, service delivery and quality management.
4. Assess the feasibility and acceptability of the renewed program for women.

The first two objectives outlined above are being undertaken through the Medical Services Advisory Committee (MSAC) review process.

MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Schedule (MBS) in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. MSAC adopts an evidence-based approach to its assessments, based on reviews of the scientific literature and other information sources, including clinical expertise.

MSAC's terms of reference and membership are listed in Appendix A. MSAC is a multidisciplinary expert body, comprising members drawn from such disciplines as: diagnostic imaging, pathology, surgery, internal medicine and general practice, clinical epidemiology, health economics, consumer health and health administration.

Both liquid-based cytology (LBC) and human papillomavirus (HPV) testing have been reviewed previously by MSAC. But it is now time for these technologies to be assessed in conjunction with a review of the screening age range and interval and potential screening pathways.

This report summarises the assessment of current evidence for cervical screening; a separate report models the cost-effectiveness of cervical screening in the Australian context ('Effectiveness modelling and economic evaluation in the Australian setting').

Background

Population-based screening

The World Health Organization defines screening as the presumptive identification of unrecognised disease or defects by means of tests, examinations or other procedures that can be applied rapidly. Screening is intended for all people, in an identified target population, who do not have symptoms of the disease or condition for which they are being screened. The process can identify: disease pre-cursors; early disease; or disease markers. Screening intends to detect disease at an early stage when it can be better treated. Screening can reduce the risk of developing or dying from a disease, but it does not guarantee that disease will not occur, or if it occurs, that it can be cured. A 'positive' screening test identifies people who are at an increased likelihood of having the condition and who require further investigation to determine whether or not they have the disease or condition (AHMAC 2008).

In population-based screening, a test is offered systematically to all individuals in a defined target population within a framework of agreed policy, protocols, quality management, monitoring and evaluation (AHMAC 2008). The target population is a clearly defined sub-group of the whole population that has been shown by strong scientific evidence to be most at risk of getting the disease and that will gain the most health benefit from early detection of the disease or its precursors.

Like all tests, screening tests are never 100% accurate. Therefore, most population-based screening programs require the target population to be screened at regular intervals. The screening intervals are determined based on an understanding of the progression of the disease and a risk/benefit analysis of the period of time that will optimise early detection and minimise over-investigation of any changes that occur in the screening interval. Three national, population-based cancer screening programs exist in Australia: BreastScreen Australia, the National Cervical Screening Program, and the National Bowel Cancer Screening Program.

The Australian Government has developed a Population Based Screening Framework (AHMAC 2008), based on the World Health Organization's principles of screening. The aim of the Framework is to provide guidance for decision-makers when considering potential population-based screening programs in Australia by outlining the requirements of population-based screening.

The Framework states that high level evidence is essential to inform screening programs, as screening has the potential both to provide benefits and cause harm to an otherwise healthy person. Some of the key issues for consideration here are potential harms such as anxiety, discomfort, adverse effects, follow-up investigations, over-diagnosis and possible treatment. These need to be carefully weighed against the potential benefits from screening in terms of reduced morbidity and mortality.

Table 1 Key terminology

Term	Definition
Incidence:	The number of new cases of cervical cancer diagnosed per 100,000 women in a year.
Morbidity:	Illness or disease resulting from cervical cancer or its precursors.
Mortality:	The number of deaths from cervical cancer per 100,000 women in a year.
Cytology:	The examination of cells from the cervix (usually collected by a Pap test) through a microscope.
Histology:	The examination of tissue from the cervix (usually collected by a biopsy) through a microscope.
Population-based screening program:	A program where screening is offered systematically to all individuals in a defined target group by applying a screening test for a disease or risk marker.
Screening test:	A comparatively simple investigation that is able to classify people according to their likelihood of having a particular disease or risk marker for a disease.

Cervical cancer in Australia

Cervical cancer affects the cells of the uterine cervix; it may arise from the squamous cells that cover the outer surface of the cervix (known as 'squamous cell carcinoma') or from glandular cells in the cervical canal (known as 'adenocarcinoma').

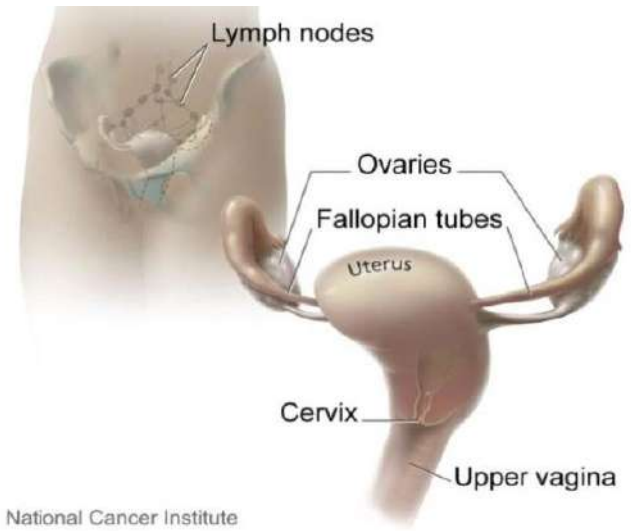


Figure 1 Anatomy: the female reproductive system (cervix, ovaries, uterus)
Source: Reproduced with permission of the National Cancer Institute

In Australia in 2008, 65.1% of cervical cancers were squamous cell carcinoma and 25.7% were adenocarcinoma, with adenosquamous (3.3%) and other cervical cancers (5.9%) making up the remainder (AIHW 2012). Cervical cancer is the thirteenth most common cancer affecting Australian women (excluding basal and squamous cell carcinoma of the skin). In 2008, the latest year for which data are available, there were 9.3 new cases of cervical cancer and 1.9 deaths per

100,000 women aged 20 to 69 years. Incidence of cervical cancer and mortality is much higher in Aboriginal and Torres Strait Islander women, with incidence at 22.3 cases and death at 9.7 per 100,000 women in the period 2003 to 2007 (AIHW 2012).

Causes of cervical cancer

The understanding of the natural history of cervical cancer has progressed significantly in recent years (AIHW 2012). It is now recognised that cervical cancer is a rare outcome of persistent infection with HPV, and that infection with a high-risk HPV type is necessary, although not sufficient, for the development of cervical cancer (Bosch et al. 2002). Most women with high-risk HPV infections will not develop cervical cancer.

Currently 15 high-risk types of HPV are recognised. HPV types 16, 18 and 45 are most predominantly associated with cervical cancer, with HPV types 16 and 18 detected in 70–80% of cases of cervical cancer in Australia (Brotherton 2008).

Infection with one or more of the 40 HPV types is extremely common, with infection rates of this sexually transmitted infection peaking in women in young adulthood (the period following sexual debut). Most HPV infection, even with high-risk types, is asymptomatic and cleared by the immune system within a year; however, in up to 10% of women the infection can persist, and in a very small number of women, persistent infection with high-risk HPV may eventually lead to cervical cancer (AIHW 2012).

Two Australian programs are designed to prevent cervical cancer: the National HPV Vaccination Program (NHPVVP) and the National Cervical Screening Program. Details of these two programs are provided below.

Prevention of cervical cancer: National HPV Vaccination Program

Australia was one of the first countries to introduce a national HPV vaccination program (NHPVVP). Two HPV vaccines are currently registered for use in Australia—*Gardasil*® and *Cervarix*® , both of which are prophylactic vaccines, which means they need to be administered prior to HPV infection. These HPV vaccines protect against high-risk HPV types 16 and 18. As noted earlier, HPV types 16 and 18 are the two main high-risk HPV types that can lead to cervical cancer; they are detected in 70–80% of cervical cancers in Australia (Brotherton 2008). Gardasil also protects against HPV types 8 and 11, which are commonly associated with genital warts in males and females. *Gardasil*® is the HPV vaccine currently used for the NHPVVP.

The NHPVVP commenced in April 2007 with the provision of free HPV vaccinations through school-based programs, to females aged between 12 and 13 years. This was complemented by a two-year catch-up program for 13 to 18-year-old females in schools, and 18 to 26-year-old females, which was delivered through general practitioners. In 2011, 71.2% of Australian females were vaccinated with three doses of HPV vaccine by 15 years.

From 2013, the current school-based program for females aged 12–13 years is being extended to offer HPV vaccination free to males aged 12–13 years, with a catch-up program in 2013 and 2014 for males aged 14–15 years.

It is expected that the HPV vaccine will decrease the number of cervical abnormalities detected and the number of cases of cervical cancer, initially in the youngest women in the screening target age group. Early data from the Victorian Cervical Cytology Registry have shown a decrease in the number of histologically confirmed high-grade cervical lesions detected in women aged 20–24

from 19.8 per 1,000 in 2008 to 15.7 per 1,000 in 2011 (VCCR 2011). This decrease is expected to continue. However, since cervical cancer incidence peaks in women over the age of 45–50 years, it will be some years before the full effects of vaccination are manifested. Male vaccination from 2013 will have an incremental effect on infection rates in women, as it will reduce the risk of unvaccinated women acquiring high-risk HPV from a male partner.

As the HPV vaccine protects only against two of the 15 high-risk HPV types, the possibility remains of cervical cancer developing due to infection with one or more of the other types. Consequently, the Department of Health and Ageing (DoHA) currently recommends that women, whether vaccinated or unvaccinated, should be screened for cervical cancer in accordance with the policy of the National Cervical Screening Program and the 2005 National Health and Medical Research Council (NHMRC) Guidelines: Screening to prevent cervical cancer: Guidelines for the management of asymptomatic women with screen detected abnormalities (the NHMRC Guidelines) (DoHA 2011).

Prevention of cervical cancer: National Cervical Screening Program

Current screening policy

Cervical screening was carried out opportunistically in Australia from the 1960s until 1991, when an organised screening program (now called the National Cervical Screening Program (NCSP)) was introduced. The NCSP is jointly funded by the Australian and state and territory governments. It currently sets out the following policy for screening:

- Routine screening with Pap tests should be carried out every two years for women who have no symptoms or history suggestive of cervical pathology.
- All women who have ever been sexually active should start having Pap tests between the ages of 18 and 20 years, or one or two years after first having sexual intercourse, whichever is later.
- Pap tests may cease at the age of 70 years for women who have had two normal Pap tests within the last five years. Women over 70 years who have never had a Pap test, or who request a Pap test, should be screened.

This policy applies only to women without symptoms that could be due to cervical pathology. Women with a past history of high-grade cervical lesions, or who are being followed up for a previous abnormal smear, are managed in accordance with the NHMRC Guidelines.

The Pap test

Cervical screening aims to detect and treat precancerous abnormalities in cervical cells before their potential progression to cervical cancer, thereby reducing cervical cancer incidence, and morbidity and mortality from this disease (AIHW 2012).

As noted above, cervical abnormalities are caused by acute infection with HPV; most will regress within a short period of time and do not require treatment. High-grade abnormalities usually occur after persistent infection with high-risk HPV. The probability of a high-grade abnormality progressing to cancer increases with age and the extent of abnormality, but cancer is still a very rare outcome (NHMRC 2005).

The screening tool currently used by the NCSP is the cytology from the Papanicolaou smear, or 'Pap test'. During a Pap test, cells are collected from the transformation zone of the cervix—the area of the cervix where the squamous cells from the outer opening of the cervix and glandular

cells from the endocervical canal meet. This is the site where most cervical abnormalities and cancers are detected. These cells are then transferred onto a slide for conventional cytology, and sent to a pathology laboratory for assessment. The cells collected are then examined under a microscope to look for abnormalities (AIHW 2012).

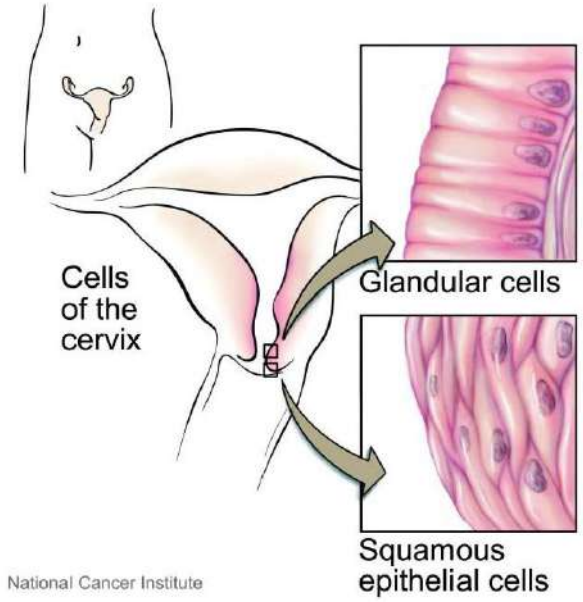


Figure 2 Anatomy: cells of the cervix
Source: Reproduced with permission of the National Cancer Institute

While cervical cytology, the examination of the cells collected from the cervix, is a very useful test, it should be emphasised that it is not a diagnostic test. This is unlike cervical histology, which is the examination of tissue collected from the cervix by means of a biopsy, in order to confirm the presence of an abnormality. As a screening test, the aim of cervical cytology is to identify those women who have a cervical abnormality, as indicated by the presence of abnormal cells in the specimen collected; women who are identified as having a cervical abnormality require further diagnostic tests. Since the Pap test only collects a sample of cells from the surface of the cervix, some judgment is required in the interpretation of sampled cells. Cervical cytology cannot accurately reveal all abnormalities that may exist in the cervical tissue in situ in a single sample (AIHW 2012).

The ability of cervical cytology to accurately exclude screenees without disease with few false positives, that is, the specificity of the screening tests is very high. Specificity estimates range from 62% to 98% (with Australia at the higher end of the scale) in an International Agency for Research on Cancer (IARC) review. However, the ability to accurately detect precancerous cervical lesions, that is, the sensitivity of a single cervical cytology test is only moderate in contrast (40–86%), indicating a greater likelihood of false negatives (IARC 2005). The strength of cervical screening comes from repeating the cervical cytology test at agreed rescreening intervals, which increases the likelihood of detecting precancerous abnormalities over the long pre-invasive stage of squamous cervical cancers (Dickinson 2002). The recognition of cervical screening as a program of

rescreening at regular intervals rather than as a single opportunistic test is an important distinction (Dickinson 2002).

Role of the cervical screening registries

Key to Australia's cervical screening program are the cervical screening registers that were established along with the cervical screening program in each state and territory. Cervical screening registers fulfil many important roles, including:

- sending reminder letters to women overdue for screening (the cervical screening registers do not currently invite women to participate in the NCSP)
- providing a safety net for women who have not had follow-up of an abnormal result
- providing cytology laboratories and cervical cytology providers with previous results for a woman, to allow a more detailed evaluation of present findings
- providing pathology laboratories with essential correlation data for cytology and histopathology for internal and external quality assurance monitoring.

State and territory cervical cytology registries also provide data on the epidemiology and natural history of precancerous lesions, as well as providing data annually to the Australian Institute of Health and Welfare (AIHW) for national monitoring of the NCSP.

Outcomes of the National Cervical Screening Program

Since the introduction of the NCSP, the incidence of cervical cancer has been significantly reduced: from 17.2 per 100,000 women in 1991 in the target age group (20–69 years) to 9.3 in 2008. The mortality rate from cervical cancer has also been significantly reduced in the target age group, from 4.0 per 100,000 women in 1991 to 1.9 in 2007 (AIHW 2012).

However, the incidence of cervical cancer in Aboriginal and Torres Strait Islander women was more than twice that of non-Indigenous women (2004–2008), and death from cervical cancer of Aboriginal and Torres Strait Islander women was five times the non-Indigenous rate (2003–2007) (AIHW 2012).

Participation rates exceed 55% and 70% within two and three year intervals respectively for women in the target age group and over 80% of the target age group have had a Pap test within five years (AIHW 2012).

Renewal of the National Cervical Screening Program

Rationale for the Renewal

The science of cancer is one of the most rapidly changing areas in health and, while the success of the NCSP cannot be disputed, the environment in which the program operates has changed. Since the introduction of the NCSP in 1991, there is a greater depth of knowledge and understanding about the natural progression of cervical abnormalities and the development of cervical cancer.

Evidence about the screening age range and interval has also changed over time, and new tests for the early detection of precancerous cervical changes have been developed. Furthermore, young Australian females and males are now provided with the opportunity to be vaccinated against HPV which will prevent the majority of, but not all, high-risk HPV infections that may lead to cervical cancer.

By international standards, the current NCSP policy is regarded as intensive and consequently, a number of national committees have recommended a review of the cervical screening policy. This includes the NHMRC's recommendation that the screening age range in Australia be reviewed when it approved its *Guidelines for the management of asymptomatic women with screen detected abnormalities* in June 2005. In 2011, the former Australian Population Health Development Principal Committee of the Australian Health Ministers' Advisory Council (AHMAC) also endorsed the plan to renew the NCSP. Thus, the Renewal of the NCSP commenced in late 2011 and is planned for completion by mid-2014.

Renewal aim and objectives

The Renewal aims to ensure the continuing success of the NCSP, and that all Australian women, HPV-vaccinated and unvaccinated, have access to a cervical screening program based on current evidence and best practice. The Renewal process will ensure the renewed program remains safe and continues to improve health outcomes of Australian women.

The objectives of the Renewal are to:

1. assess the evidence for screening tests and pathways, the screening interval, age range and commencement for both vaccinated and non-vaccinated women

2. determine a cost-effective screening pathway and program model

3. investigate options for improved national data collection systems and registry functions to enable policy, planning, service delivery and quality management

4. assess the feasibility and acceptability of the renewed program for women.

The first two objectives outlined above are being undertaken through the Medical Services Advisory Committee (MSAC) review process.

A Decision Analytic Protocol (DAP) was commissioned by MSAC and ratified by the Protocol Advisory Sub-committee in September 2012. The DAP is multi-tiered, and outlines a range of possible screening tests and screening pathways, in the context of the availability of the HPV vaccination. Table 10, in the 'Methodology' section, provides an overview of the cervical screening scenarios that will be compared to the current NCSP.

HPV vaccination and the Renewal

The impact of HPV vaccination will be an important consideration when assessing the most appropriate cervical screening pathway for the future, especially for younger women. As noted previously, the introduction of HPV vaccination:

- will provide protection for young women against the two most common forms of high-risk HPV, that is, strains 16 and 18

- will reduce the number of cervical abnormalities detected, impacting on the effectiveness and cost-effectiveness of the screening test

- in the long term, will reduce the incidence of cervical cancer; this may also reduce the effectiveness and cost-effectiveness of the current NCSP pathway.

These impacts need to be considered to ensure that both the NCSP and the National HPV Vaccination Program are safe, efficient and effective.

Underscreened women and the Renewal

The Renewal will also address the need to reach unscreened and underscreened women, in order to further reduce deaths from cervical cancer in Australia. Towards that aim, the Renewal will be considering self-collection for HPV tests (that require women to take a sample of cells directly from the cervix) for unscreened or underscreened women only. Additionally the Renewal will also consider the introduction of an invitation and recall system (rather than the current overdue reminder system), in order to increase participation in the screening program.

Framework for the evaluation

Cervical cancer trends in Australia

Overall incidence and mortality

The aim of both the NHPVVP and the NCSP is to reduce the incidence of and mortality from cervical cancer. The overall incidence (the number of new cases per 100,000 women per year) of cervical cancer in Australia has decreased since 1982 with a dramatic decline over the 1990s following the commencement of the NCSP in 1991 (Figure 3).

In 1982, 14.2 new cases of cancer occurred per 100,000 Australian women. In 1992, this number decreased to 12.2 new cases, and in 2002 to 6.8 new cases per 100,000 Australian women. The number of new cases of cervical cancer has held relatively steady since then; in 2009 (the most recent year for which the incidence data are available), cervical cancer incidence among Australian women was 6.8 per 100,000 Australian women (AIHW 2012).

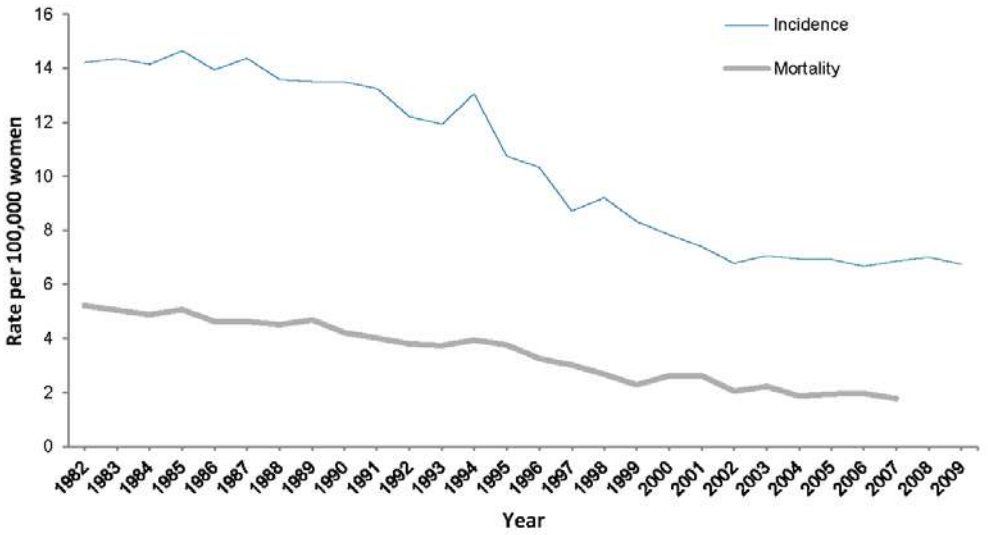


Figure 3 Overall incidence and mortality from cervical cancer in Australia

Overall mortality from cervical cancer has been steadily decreasing in Australia. In 1982, the mortality rate was 5.2 per 100,000 Australian women; in 1992, the rate diminished to 3.8, and in 2002 to 2.0. In 2007 (the most recent year for which data are available), 1.8 women per 100,000 died of cervical cancer.

Age-stratified incidence and mortality

The incidence of cervical cancer has declined across all age groups since the program commenced in 1991; however, this impact has been greatest in older women (Figure 4). For women aged 20–24 years, cervical cancer is very rare and the magnitude of difference in incidence in absolute terms is very small, with cancer registry data showing a five-year age-specific cervical cancer incidence in the pre-screening period 1983–87 of 2.7 per 100,000 compared to 1.4 per 100,000 in 2002–07. Cervical cancer incidence rates increase until the age of 30 and then plateau with rates of 10.5

cases per 100,000 in the 30–34 age group and 10.1 in the 50–54 age group. In the above-65 age group, the incidence rates remain similar to those in the target age range ranging from 10.0 cases per 100,000 (70–74 year olds) to 13.6 cases per 100,000 (80–84 year olds) in 2005–2009 (AIHW 2012).

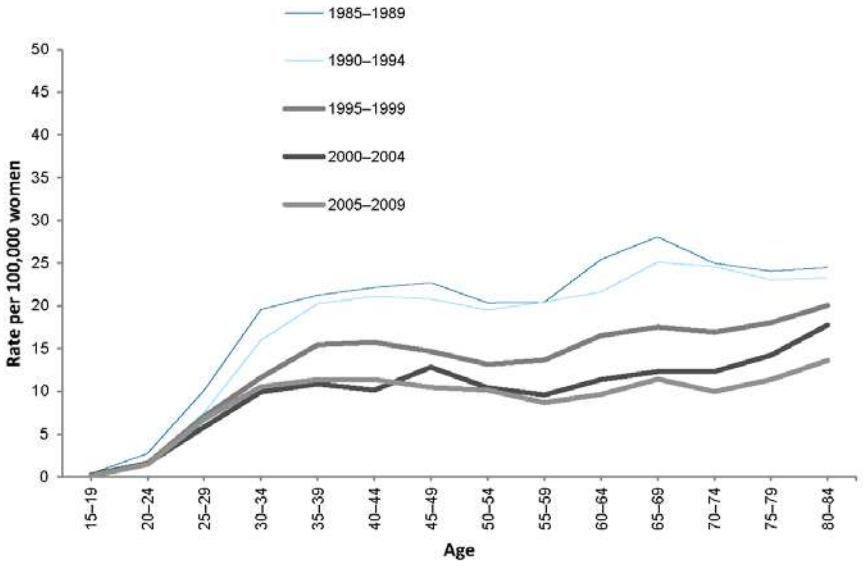


Figure 4 Age-stratified incidence of cervical cancer in Australia, 1985–2009

The rate of mortality from cervical cancer in 2003–2007 was approximately one-half or less the rate of mortality from cervical cancer in 1983–1987 for all age groups with the impact again greater in older women (Figure 5). Among the under-25s, the mortality rate in 2003–07 was negligible (0.0 and 0.1 among 15–19-year-olds and 20–24-year-olds respectively) as it was in 1985–1989 (0.0 and 0.2 among 15–19-year-olds and 20–24-year-olds respectively). In 2003–2007, the mortality from cervical cancer in the age range of 25–64 ranged from 0.5 per 100,000 (in the 25–29 age group) to 4.8 (in the 60–64 age group). In the above-65 age group, mortality ranged from 4.4 per 100,000 (65–69 year olds) to 9.7 (80–84 year olds) (AIHW 2012).

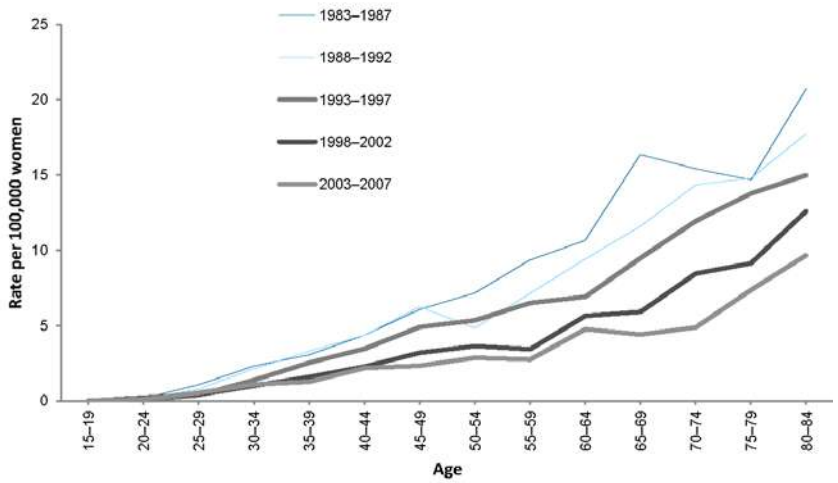


Figure 5 Age-stratified mortality from cervical cancer in Australia, 1983–2007

In Canada, the impact of the introduction of cervical cancer screening in the 1970s is markedly similar to that in Australia (Canadian Task Force on Preventive Health Care 2013). In particular, the incidence and mortality rates in women aged 20–24 are similarly low, with an incidence rate of 2.7 per 100,000 and a mortality rate of 0.1 per 100,000 in 1972–1976 (Canadian Task Force on Preventive Health Care 2013; Dickinson 2013). In 2002–2007 there was a mortality rate of 0.2 per 100,000 (indicating no reduction in mortality over this 20-year period) and an incidence rate of 1.3 per 100,000 (a potential reduction in incidence of 1.4 per 100,000).

In England, the incidence of cervical cancer has also fallen over the period 1982–2006 since the introduction of a cervical cancer screening program in 1988 (Foley et al. 2011). However, a recent study has shown that while incidence rates have declined continuously in women aged 40–79, the incidence in women aged 20–29 has increased over the period 1992–2006 (Foley et al. 2011). Incidence in women of this age born in 1977–1981 has significantly increased compared with the three previous five-year cohorts. The women born in 1977–1981 would have received at least two invitations to screen before the start age for screening in England was raised to 25 years in 2005, thus this change in screening policy cannot explain these data. In Australia, with the introduction of the NHPVVP, it is expected that the incidence of cervical cancer in women aged under 30 years will decline in the future. Early data from the Victorian Cervical Cytology Registry have already shown a decrease in the number of histologically confirmed high-grade cervical lesions over the period 2008 to 2011 (VCCR 2011).

Population targeted by the screening program

Age range and frequency of cervical screening

Under the current NCSP, it is recommended that screening should commence for all sexually active women between the ages of 18 and 20, or one to two years after first having sexual intercourse, whichever is later; and that two-yearly screening should cease at age 70. It has been suggested by the applicant that IARC recommendations should be considered. In 2005, IARC recommended that cervical screening should cover women aged 25 to 65 and that women should undergo screening once every three years up to the age 49, and every five years thereafter (Table 2).

Table 2 Comparison of the current National Cervical Screening Program (NCSP) age range and screening interval, and recommendations from the International Agency for Research on Cancer (IARC) 2005

	NCSP, current in 2013	IARC, 2005
Primary screening test	Conventional cytology	Conventional cytology or LBC if funds available and local feasibility If HPV adopted, not recommended before 30 years of age
Age range	-	-
Start age	18–20 years or 1–2 years after becoming sexual active	25 years
Exit age	70 years if previous 2 screens within last 5 years both normal	65 years if previous 2 screens within the last 10 years both normal
Screening interval	-	-
<25 yrs	2-yearly	Nil
25–50 yrs	2-yearly	3-yearly
51–65 yrs	2-yearly	5-yearly

	NCSP, current in 2013	IARC, 2005
66–70 yrs	2-yearly	Nil

A recent United Kingdom (UK) study (Sasieni et al. 2009) assessed the effectiveness of cervical screening by age and concluded that cervical screening in women aged 20 to 24 has little or no impact on rates of invasive cervical cancer up to age 30; however, some uncertainty exists regarding its impact on advanced-stage tumours in women under age 30. In contrast, screening older women leads to a substantial reduction in incidence of and mortality from cervical cancer. Available information (see Appendix C) indicates that, with the exception of the United States (US) and some European countries, the commencement age for screening is generally around 25 years of age.

The applicants have indicated that the age range for cervical screening should be reviewed in this assessment. While the current screening program recommends that cervical screening may cease at the age of 70 years for women who have had two normal cytology tests within the last five years, it should be recognised that the lifetime impact of regular cervical screening will benefit older women past the exit age for the program. Therefore, it will be important to ensure the assessment captures this positive effect on older women outside the target population.

Cervical screening programs in developed countries vary in their recommendations for the age range (start and exit age) and frequency of screening (the recommended interval between screening episodes), as summarised in Appendix C. Screening start ages vary between 15 and 30 years, exit ages vary between 60 and 70 years, and screening intervals vary from 1 to 5 years. These differences in age range and screening intervals translate to a large difference in the total number of screening tests a woman will receive in her lifetime in different countries. For example, women in the Netherlands can expect to receive 6 to 8 cytology screening tests compared to over 50 tests in Germany.

Comparing recommended cervical screening age ranges between Australia, Canada, New Zealand, the UK and the US, only New Zealand uses the same screening age range of 20–69 years as Australia; other countries recommend screening for a narrower age range (Appendix C).

In 2013, most countries that use cytology as the primary screening test (LBC or conventional Pap test) also use a longer screening interval than Australia, with screening every three years currently recommended in Canada, New Zealand and the US. The UK follows the IARC (2005) recommendations for three-yearly screening for younger women and five-yearly screening for older women (based on evidence reported by Sasieni et al. (2003)). In the US, the United States Preventive Services Task Force (USPSTF) recommends three-yearly cytology screening. However, as an alternative for women aged 30 to 65 years who want to lengthen the screening interval, the USPSTF recommends 'co-testing' with both cytology and HPV every five years (Moyer 2012).

The applicants have indicated that the present review of the NCSP age range and screening interval should particularly consider the impact of following the IARC (2005) recommendations for a narrower age range and longer screening interval using cytology as the primary screening test.

The impact of moving to a longer screening interval or younger exit age (after a negative test result on the previous screening episode) will depend on the type of test used for primary screening. Estimates based on using cytology as the primary screening test will not apply to screening strategies that use HPV as the primary screening test. This is because a negative HPV test and a

negative cytology test are associated with different risks of cervical cancer over time (Kitchener 2011b).

The epidemiological principles for assessing evidence about the safety and effectiveness of changing the screening age range and interval are described in the 'Methodological issues' section on page 29.

Program participation: Call–recall versus reminder systems

The age-standardised two-yearly participation rate of the target population (women aged 20–69 with an intact cervix) in 2010–2011 was 57.2% (AIHW 2013). Participation in the national cervical screening program peaks in women aged 45–49 at 63.0%, followed by women aged 50–54 at 62.6% (Table 3). Participation is low, and declining, in women aged 20–24 (falling from 45.3% in 2008–2009 to 42.6% in 2010–2011). In 2010–2011, 2.7% of women screened were under 20, and 1.3% were aged 70 or over (AIHW 2013).

Table 3 Participation by age, 2010–2011

Age group	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69
Women	334,336	419,220	435,404	464,527	444,963	425,158	378,138	311,591	261,433	166,428
Crude rate	42.6	51.9	58.1	60.3	61.6	63.0	62.6	60.3	57.7	50.4

Note: Crude rate is the number of women screened in 2010–2011 as a percentage of the Australian Bureau of Statistics (ABS) estimated resident population for women aged 20–69, adjusted to include only women with an intact cervix using age-specific hysterectomy fractions derived from the AIHW National Hospital Morbidity Database.
Source: AIHW analysis of state and territory cervical cytology register data.

Participation was similar across states and territories and remoteness areas but showed a clear trend of increasing participation with increasing socioeconomic status from 52% in areas of lowest socioeconomic status to 63% in areas of highest socioeconomic status (AIHW 2013). Participation in cervical screening cannot be measured nationally for Aboriginal and Torres Strait Islander women; however, there is evidence that these women are underscreened (Coory et al. 2002; Binns & Condon 2006).

It is among the unscreened and underscreened population that the vast majority of invasive cervical cancers occur; the Victorian Cervical Cytology Registry estimates that over 80% of Victorian women diagnosed with invasive cervical cancer in 2009 have either never been screened or were lapsed screeners prior to their cancer diagnosis (VCCR 2011). Reliable national data on cervical cancer incidence in Aboriginal and Torres Strait Islander women are not available, but data from four states and territories (New South Wales, Queensland, Western Australia and the Northern Territory) do enable analysis and include around 85% of Aboriginal and Torres Strait Islander people. Over the five-year period 2004–2008, Aboriginal and Torres Strait Islander women in these four jurisdictions had a significantly higher incidence of cervical cancer (22.3 new cases per 100,000 women) compared with non-Indigenous women (8.5 new cases per 100,000 women) (AIHW 2013). Reaching underscreened and unscreened women is a key measure to further reducing deaths from cervical cancer in Australia.

In England in 2008–2009, the cervical cancer diagnosis and death of a female celebrity (reality TV star Jade Goody) resulted in an increase in attendance for cervical cancer screening in England and Wales (Lancucki et al. 2012; Marlow et al. 2012). Similarly, a spike in Australian cervical cancer incidence data can be seen in 1994–1995 (Figure 3), corresponding to an increase in attendance for screening following the death of a 29-year-old Australian women (Rhonda O'Shea). There was significant media coverage of this case in Australia, increasing public awareness and hence attendance for screening. In the UK in 2008–2009, the increase in attendance due to the 'Jade

Goody effect' was in a greater proportion of women whose scheduled screening was overdue, rather than overscreening (Lancucki et al. 2012). The outcomes of the estimated 501,000 additional attendances showed a similar distribution to usual attendances, with 370 cases of suspected neoplasia. The influence of this case in increasing attendance in England was greater in women aged 25 to 49 years (Lancucki et al, 2012) and in women from lower socioeconomic backgrounds or with fewer educational qualifications (Marlow et al, 2012).

The Renewal includes the consideration of an invitation and recall system (Table 10) as a measure to support high rates of participation in the screening program. Evidence for invitation and recall systems was not systematically appraised in this review but a brief overview is provided below. A modelled analysis of cervical cancer screening, diagnosis and treatment using an invitation and recall system in HPV-vaccinated and unvaccinated populations in Australia is in progress to explore the potential long-term benefits and harms of these alternative screening strategies in the Australian setting.

Currently, each Australian state and territory operates a Pap Test Register (PTR), which is a database of Pap tests and related follow-up tests for women residing in that state or territory. One of the key functions of each PTR is to send out a reminder letter to a woman, if more than 27 –36 months (depending on state/territory) have elapsed since her previous Pap test (Canfell et al. 2006). In contrast to this reminder system, in the UK women receive both an invitation before their rescreen is due and a reminder letter 6 to 12 months later with a screening interval of three years. In a comparison between the two programs (using New South Wales (NSW) data), it was found that in the UK, the majority of rescreening happens around three years after a negative screen whereas in NSW the timing of rescreening is much more variable with rescreening most likely to occur either at about one year or after two years (Canfell et al. 2006).

A literature review for the Victorian Cervical Cytology Register identified four studies reporting on the effectiveness of mailed invitations and reminders for cervical screening; three of these demonstrated a statistically significant positive impact on the uptake of screening and one showed no improvement. The overall conclusion, across breast, cervical and bowel screening was that invitations and reminders do increase uptake, at least among mid to high-socioeconomic groups (Day et al. 2010). A Cochrane review of interventions to increase the uptake of cervical screening was published in 2011 (Everett et al. 2011). Seventeen RCTs or clustered RCTs were identified evaluating the effectiveness of invitation letters and a meta-analysis of 12 trials was undertaken assessing 99,651 participants which found that women who received invitation letters to attend cervical screening programs had a significantly higher uptake of screening than women who received usual care or no invitation (RR 1.44 (1.24–1.52)), although there was substantial heterogeneity between studies (I^2 = 72%). The authors concluded that there is evidence to support the use of invitation letters to increase the uptake of cervical screening.

Screening technologies

Liquid-based cytology

Liquid-based cytology (LBC) is a method of preparing cervical samples for examination in the laboratory. The sample is collected in a similar way to conventional cytology; however, the head of the spatula, broom or brush containing the cells is broken or rinsed into a vial containing preservative liquid. The sample is then sent to the laboratory for processing to remove obscuring material such as mucous, pus or blood before being placed on a slide.

There are currently two LBC systems available in Australia. These two systems use different technical methods for processing the cells before placement on a slide. One uses a cell filtration

system (ThinPrep® Pap system, Hologic [Australia] Pty Ltd) and the other a cell enrichment system (SurePath™ LBC system, Beckton Dickinson Pty Ltd). Automated image analysis can also be used with LBC, which allows the cytotechnologist to be directed to an area on the slide that is most likely to contain abnormal cells. Automated image analysis aims to reduce the time required to read a slide and reduce detection error.

Previous assessments of LBC by MSAC

MSAC has previously considered LBC and automated image analysis on a number of occasions. Details of these assessments are provided in Appendix J. The 2002 assessment found that there was insufficient evidence to determine whether LBC was equal or superior in effectiveness compared with conventional cytology. The model used indicated that LBC was associated with greater costs per woman than conventional cytology. Since there was insufficient evidence to support a claim that LBC is superior to conventional cytology in detecting high-grade lesions or invasive cancer, there was no evidence to suggest that LBC would be cost-effective at the proposed price and MSAC advised there was insufficient evidence to support public funding of LBC for cervical screening.

In 2003, MSAC considered the safety, effectiveness and cost-effectiveness of automated image analysis for cervical screening cytology compared with manual processing (see Appendix J for details of this assessment). MSAC determined there was insufficient evidence to assess whether automated image analysis is as effective as manual processing for cervical screening cytology. Given the lack of clinical evidence, an economic evaluation was not conducted and MSAC advised that there was insufficient evidence to support public funding of automated image analysis for cervical screening.

The 2009 assessment considered LBC using automated image analysis systems as well as manual LBC. The available evidence demonstrated that compared to conventional cytology, manual LBC provided no statistically significant increase in sensitivity or specificity. Automated LBC detected more cervical intraepithelial neoplasia (CIN) 2+ lesions compared to conventional cytology, but results from one trial raised uncertainty about whether this difference is attributable to LBC alone, to the automation-assisted reading system or a combination of both. A modelled analysis found that automated LBC would be associated with a cost of $194,835 per life year saved (LYS). Manual LBC was associated with a cost of $126,315 per LYS to $385,982 per LYS, depending on the level of reimbursement. MSAC concluded LBC is at least as effective as conventional cytology, but is not cost-effective at the price requested and should not be supported for public funding.

In 2013 MSAC is considering evidence on manual cell enrichment LBC. As such MSAC will consider evidence on cell enrichment LBC in respect to cervical cancer screening. The applicants have indicated that re-assessment of LBC may be beneficial alongside a review of the cervical screening age range and interval and HPV testing.

HPV testing

There are over 100 different types of HPV. Some of these collectively referred to as 'oncogenic types', have been linked to the development of cervical abnormalities and cervical cancer. HPV tests detect the genetic material of high-risk oncogenic types of HPV associated with cervical cancer. There are different methods available for HPV testing: the two most common are polymerase chain reaction (PCR) and hybrid capture which utilises RNA probes to hybridise to the viral DNA.

The most commonly used test is the Hybrid Capture 2 (HC2) test (Qiagen Inc.) which identifies 13 high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) and five low-risk HPV types. A positive result indicates infection with one of more subtypes but cannot be used to identify the subtype.

Currently, in Australia HPV testing is generally undertaken using a pooled method that tests for a number of different oncogenic types, to give an overall result of positive or negative. However, increasingly, clinically positioned HPV tests which provide a partial genotyping facility to signal the presence of the most common types implicated in cancer (such as HPV 16/18), as distinguished from grouped 'other oncogenic' HPV types, are becoming available (for example the Cobas® 4800 HPV test (Roche)).

Previous assessments of HPV testing by MSAC

HPV DNA testing as a triage tool

In 2002, MSAC considered the use of HPV DNA testing as a triage tool. Full details of the assessment are provided in Appendix J. The evidence provided indicated that HPV DNA testing was more sensitive but less specific than cytology, although the evidence did not support widespread implementation. The assessment concluded that additional high-quality studies using an acceptable reference standard, such as histological confirmation of cytology results, would be useful in allowing a valid and reliable judgment of the sensitivity and specificity of HPV DNA testing. A decision analytic model indicated that HPV DNA testing was both more expensive and less effective in detecting high-grade lesions than the management plan currently recommended by the NHMRC, but the model was particularly sensitive to the estimated prevalence of high-grade lesions in women. MSAC advised that there was currently insufficient evidence to support public funding at the time for the use of the HPV DNA test for triaging of women with equivocal cervical screening results.

In 2009, HPV DNA testing as a triage tool was re-assessed. Details of this assessment are provided in Appendix J. Comparative accuracy studies provided strong evidence that an immediate HPV DNA triage test is a more sensitive test than a single repeat cytology test for detecting CIN2+ lesions in women with possible low-grade squamous intraepithelial lesion (pLSIL), and has similar specificity to cytology. An immediate HC2 HPV triage test is no more sensitive than a single repeat cytology test for detecting CIN2+ lesions in women with LSIL, and has lower specificity, but colposcopy referral rates may be favourable compared with a strategy of two annual repeat cytology tests. Restricting the HPV DNA triage test to older age groups is associated with higher specificity and a lower colposcopy referral rate, but a smaller gain in sensitivity, compared with its use in all age groups. A modelled analysis predicted that, compared with current practice, a strategy of performing the HPV DNA triage test for women aged 30+ years produces an incremental cost-effectiveness ratio (ICER) of $75,739 per life year saved (LYS) if conventional cytology is used with co-collection for HPV DNA testing; or $83,496 per LYS using manual LBC with reflex HPV DNA testing; or $170,209 per LYS using automated LBC with reflex HPV DNA testing. On the basis of the available results, MSAC advised that HPV DNA triage testing in cervical cancer was not cost-effective in the Australian setting at the current price of HPV DNA testing and MSAC did not support public funding.

HPV DNA testing as a primary screening test

In 2003, MSAC considered the use of HPV DNA testing for cervical screening as either a stand-alone screening test or combined with screening by cytology. Details of the assessment are provided in Appendix J. MSAC found that there was insufficient evidence that HPV DNA testing is effective in detecting high-grade cervical lesions when used as either a stand-alone screening test

or combined with screening by cytology. Due to the lack of clinical evidence, an economic evaluation was not conducted.

The applicants have indicated that this technology should be re-assessed in conjunction with the consideration of changes in screening age and interval. There is current evidence available addressing the use of HPV DNA testing, in particular several international randomised controlled trials (RCTs) of primary HPV DNA screening compared to cytology have now reported several years of follow-up over two or three rounds of screening (Kitchener et al. 2011b; Ronco et al. 2010; Mayrand et al. 2007; Rijkaart et al. 2012; Naucler et al. 2009).

HPV testing – practitioner versus self-collected samples

HPV testing using self-collected samples has been studied as a strategy to reach underscreened populations. In self-collected samples, women can collect a sample of their own cells, usually using a tampon, cotton-tipped swab or brush, for HPV testing. Self-collection may be undertaken within the healthcare setting or outside of it (such as at home). The self-collected sample is sent to a laboratory for processing and testing. Self-collection within a healthcare setting is widely used to test for chlamydia and gonorrhoea infection.

The previous assessments of HPV testing by MSAC have not considered the use of practitioner-collected versus self-collected samples. The applicants have indicated it may be useful to consider self-collected samples for HPV testing for underscreened and unscreened women, to supplement the organised screening program using practitioner-collected samples for HPV testing.

Comparator

The comparator for any alternative cervical screening pathway (primary questions) is the current pathway promoted by the NCSP. All women between the ages of 18 (or one to two years after first having sexual intercourse, whichever is later) and 70 years of age should be screened every two years using conventional cytology.

The comparators for secondary questions are the screening strategies considered in the primary question.

Regulatory and reimbursement issues

Regulatory status of the new technologies

The Therapeutic Goods Administration (TGA) provides the regulatory framework for in-vitro diagnostic (IVD) medical devices. These include 'pathology tests and associated instrumentation used to carry out testing on human samples, where the results are intended to assist in clinical diagnosis or in making decision concerning clinical management' (TGA 2011). The framework came into effect on 1 July 2010. All IVDs supplied prior to 1 July 2010 are provided with a four year transition period (i.e. until 30 June 2014) to be brought into the regulatory framework. It would be expected that all products assessed and used as part of the NCSP would comply with the new regulatory framework. Manufacturers will be required to provide evidence that their product complies with the framework for their product to be claimed through the MBS.

Marketing status of the new technologies

Liquid Based Cytology

LBC tests with manual or automated slide reading are in vitro diagnostic tests that are not of human origin and are therefore exempt from the regulatory requirements of the *Therapeutic Goods Act 1989*. These tests comply with US Food and Drug Administration (FDA) requirements in the US and have received CE marking in Europe.

There are many currently marketed LBC preparation systems available. These systems use a variety of technical methods for storing and preparing the cervical cytology sample, some of which are patented. The ThinPrep Pap system (Cytyc Pty Ltd) includes a filtration method with membrane transfer to the slide. The SurePath and PrepStain system (Becton Dickinson Pty Ltd) uses a centrifugation system. The manufacturers assert that these technical differences translate into different performance characteristics for different LBC preparation systems.

The only currently marketed fully integrated systems of LBC and automation-assisted screening are the ThinPrep® Pap Test and Imaging System (Cytyc Pty Ltd), and FocalPoint® automated screening of SurePath cytology (Becton Dickinson Pty Ltd, formerly TriPath Imaging Inc). The FocalPoint system can be used for automated screening of conventional or LBC slides. FocalPoint was previously marketed as AutoPap, an automated screening system for conventional Pap tests. Previously marketed systems used for automated screening for quality control include AutoCyte Screen and PapNet. These systems are no longer commercially available.

HPV

A number of HPV tests are currently being used in Australia, but there is only one entry in the Australian Register of Therapeutic Goods (ARTG); this is the cobas® 4800 Human Papillomavirus (HPV) Test, sponsored by Roche Diagnostics Australia Pty Limited. The test specifically identifies HPV types 16 and 18 while concurrently detecting the rest of the high-risk types (31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68) at clinically relevant infection levels. Specimens are limited to cervical cells collected in cobas® PCR Cell Collection Media (Roche Molecular Systems, Inc.), PreservCyt® Solution (Cytyc Corp.) and SurePath® Preservative Fluid (BD Diagnostics-TriPath) (ARTG entry 187190).

Other HPV tests currently used in Australia without ARTG listing (ie under TGA transitional arrangements, as outlined above) include: HC2 test (Digene) and Cervista (Hologic). Older tests (such as Amplicor and Linear Array) are also used, although predominantly in a research setting, and the manufacturer (Roche) has no plans to submit a TGA application.

An international group has proposed guidelines stating that any new high-risk HPV test should reach a minimum relative sensitivity of at least 0.90 and a relative specificity of at least 0.98, using HC2 as a comparator test and CIN2+ as the threshold for disease. Moreover the new test should be highly reproducible (agreement >87%, minimum 500 samples) (Meijer 2009).

Current reimbursement arrangements

Conventional cytology (Pap test)

Conventional cervical cytology (or Pap test) is primarily provided through general practice and other primary healthcare settings with rebates available from the Medicare Benefits Schedule (MBS). The current MBS item descriptor for conventional cervical cytology (for screening and non-screening purposes) is outlined in Table 4.

Table 4 Current MBS Item descriptors for cervical cytology

Current MBS Item descriptors for cervical cytology
Item number 73053 Cytology of a smear from cervix where the smear is prepared by direct application of the specimen to a slide, excluding the use of liquid based slide preparation techniques, and the stained smear is microscopically examined by or on behalf of a pathologist - each examination (a) for the detection of precancerous or cancerous changes in women with no symptoms, signs or recent history suggestive of cervical neoplasia, or (b) if a further specimen is taken due to an unsatisfactory smear taken for the purposes of paragraph (a); or (c) if there is inadequate information provided to use item 73055; Fee: $19.60 Benefit: 75% = $14.70 85% = $16.70 (See para P16.11 of explanatory notes to this Category)
Item number 73055 Cytology of a smear from cervix, not associated with item 73053, where the smear is prepared by direct application of the specimen to a slide, excluding the use of liquid based slide preparation techniques, and the stained smear is microscopically examined by or on behalf of a pathologist - each test (a) for the management of previously detected abnormalities including precancerous or cancerous conditions; or (b) for the investigation of women with symptoms, signs or recent history suggestive of cervical neoplasia; Fee: $19.60 Benefit: 75% = $14.70 85% = $16.70 (See para P16.11 of explanatory notes to this Category)
Item number 73057 Cytology of smears from vagina, not associated with item 73053 or 73055 and not to monitor hormone replacement therapy, where the smear is prepared by direct application of the specimen to a slide, excluding the use of liquid based slide preparation techniques, and the stained smear is microscopically examined by or on behalf of a pathologist - each test Fee: $19.60 Benefit: 75% = $14.70 85% = $16.70 (See para P16.11 of explanatory notes to this Category)
Explanatory notes P16.11 Cervical and Vaginal Cytology - (Items 73053 to 73057) Item 73053 applies to the cytological examination of cervical smears collected from women with no symptoms, signs or recent history suggestive of cervical neoplasia as part of routine, biennial examination for the detection of pre-cancerous or cancerous changes. This item also applies to smears repeated due to an unsatisfactory routine smear, or if there is inadequate information provided to use item 73055. Cytological examinations carried out under item 73053 should be in accordance with the agreed National Policy on Screening for the Prevention of Cervical Cancer. This policy provides for: i. an examination interval of two years for women who have no symptoms or history suggestive of abnormal cervical cytology, commencing between the ages of 18 to 20 years, or one to two years after first sexual intercourse, whichever is later; and ii. cessation of cervical smears at 70 years for women who have had two normal results within the last five years. Women over 70 who have never been examined, or who request a cervical smear, should be examined. This policy has been endorsed by the Royal Australian College of General Practitioners, the Royal Australian College of Obstetricians and Gynaecologists, The Royal College of Pathologists of Australasia, the Australian Cancer Society and the National Health and Medical Research Council. The *Health Insurance Act 1973* excludes payment of Medicare benefits for health screening services except where Ministerial directions have been issued to enable benefits to be paid, such as the Papanicolaou test. As there is now an established policy which has the support of the relevant professional bodies, routine screening in accordance with the policy will be regarded as good medical practice. The screening policy will not be used as a basis for determining eligibility for benefits. However, the policy will be used as a guide for reviewing practitioner profiles. Item 73055 applies to cervical cytological examinations where the smear has been collected for the purpose of management, follow up or investigation of a previous abnormal cytology report, or collected from women with symptoms, signs or recent history suggestive of abnormal cervical cytology. Items 73057 applies to all vaginal cytological examinations, whether for a routine examination or for the follow up or

Current MBS Item descriptors for cervical cytology
management of a previously detected abnormal smear. For cervical smears, treating practitioners are asked to clearly identify on the request form to the pathologist, by item number, if the smear has been taken as a routine examination or for the management of a previously detected abnormality.

The Practice Incentives Program (PIP) Cervical Screening Incentive aims to encourage general practitioners and/or non-specialist medical practitioners to screen under-screened women for cervical cancer (Australian Government, Department of Human Services, 2012). To be eligible for the Cervical Screening Incentive practices must be registered with the PIP, the practitioner must then use one of the specific cervical screening MBS item numbers when performing a cervical smear on an under-screened women in the target age range. These item numbers indicate that the requirements have been met and will then trigger a payment to the practices, the explanatory note to these items is shown in Table 5.

Table 5 Explanatory note to MBS item numbers 2497 - 2509 and 2598 – 2616 for taking a cervical smear from a person who is unscreened or significantly under-screened

Explanatory note to MBS item numbers 2497 - 2509 and 2598 – 2616 for taking a cervical smear from a person who is unscreened or significantly under-screened
A43 Taking a Cervical Smear from a Person who is Unscreened or Significantly Under-screened - (Items 2497 - 2509 and 2598 - 2616) The item numbers 2497, 2501, 2503, 2504, 2506, 2507, 2509, 2598, 2600, 2603, 2606, 2610, 2613 and 2616 should be used in place of the usual attendance item where as part of a consultation, a cervical smear is taken from a person between the ages of 20 and 69 years inclusive who has not had a cervical smear in the last four years. The items apply only to a person between the ages of 20 and 69 years inclusive who has a cervix, has had intercourse and has not had a cervical smear in the last four years. When providing this service, the doctor must satisfy themselves that the person has not had a cervical smear in the last four years by: (a) asking the person if they can remember having a cervical screen in the last four years; and (b) checking their own practice's medical records. If significant uncertainty still remains, the doctor may also contact the state cervical screening register. A person from the following groups are more likely than the general population to be unscreened or significantly underscreened - low socioeconomic status, culturally and linguistically diverse backgrounds, Indigenous communities, rural and remote areas and older people. Vault smears are not eligible for items 2497 - 2509 and 2598 - 2616. In addition to attracting a Medicare rebate, the use of these items will initiate a Cervical Screening SIP through the PIP. A PIP Cervical Screening SIP is available for taking a cervical screen from a person who has not been screened in the last for four years. The SIP will be paid to the medical practitioner who provided the service if the service was provided in a general practice participating in the PIP Cervical Screening Incentive. A further PIP Cervical Screening Incentive payment is paid to practices which reach target levels of cervical screening for their patients aged 20-69 years inclusive. More detailed information on the PIP Cervical Screening Incentive is available from the Medicare Australia PIP enquiry line on 1800 222 032 or from the Department of Human Services website. Related Items: 2497, 2501, 2503, 2504, 2506, 2507, 2509, 2598, 2600, 2603, 2606, 2610, 2613, 2616

Liquid-based cytology

LBC by any method is not listed on the MBS. It is, in fact, explicitly excluded from the MBS (see MBS descriptors above). However, LBC is currently provided by all private pathology laboratories

for a fee additional to the MBS fee for conventional Pap tests, and is collected using the split-sample technique in conjunction with conventional Pap tests. The additional fee is paid by the patient (typically around $30 or more).

The exception to this is in Queensland (ie far north Queensland) where thin layer technology (ThinPrep®) is offered as an adjunctive test to conventional Pap tests in women meeting specific criteria (Queensland Cervical Screening Program 2008); this program is funded by the Queensland state government.

HPV

The NHMRC Guidelines recommend HPV testing for women who have undergone treatment for high-grade cervical abnormalities to monitor the effectiveness of treatment. An MBS item is available for this purpose. HPV testing for any other purpose is not currently recommended by the NCSP and a Medicare rebate is not available. The current MBS item descriptor for HPV testing (for follow-up of treated high-grade abnormalities) is outlined in Table 6.

Table 6 Current MBS Item descriptors for HPV testing

Current MBS Item descriptors for HPV testing
Item number 69418 A test for high risk human papillomaviruses (HPV) in a patient who: - has received excisional or ablative treatment for high grade squamous intraepithelial lesions (HSIL) of the cervix within the last two years; or - who within the last two years has had a positive HPV test after excisional or ablative treatment for HSIL of the cervix; or - is already undergoing annual cytological review for the follow-up of a previously treated HSIL. - to a maximum of 2 of this item in a 24 month period (Item is subject to rule 25) Fee: $64.00 Benefit: 75% = $48.00 85% = $54.40

Outcomes assessed by the Renewal

Benefits of screening

Assessment of the safety, effectiveness and cost-effectiveness of different cervical screening strategies includes an evaluation of the benefits for reducing cervical cancer incidence and mortality, and the risk of harm. The European Guidelines for Quality Assurance in Cervical Cancer Screening recognise the following outcome measures for assessment of the benefits of a cervical screening strategy, ranked in order of level of clinical importance (Arbyn et al. 2009):

1. Reduction of all-cause mortality
2. Reduction of mortality from cervical cancer: (quality-adjusted) life-years gained
3. Reduction of morbidity due to cervical cancer: incidence of invasive cancer
4. Reduction of incidence of cancer (including microinvasive cancer)
5. Reduction of incidence of CIN3 or worse (CIN3+)
6. Increased detection rate of either CIN3+ or CIN2+
7. Increased test positivity with increased, similar, or reduced positive predictive value.

Each scenario considered in the present assessment systematically reviews the evidence pertaining to the benefits as listed in this hierarchy.

Harms of screening

A number of harms may also arise from a cervical screening strategy. Potential harms raised in the literature are outlined below.

Physical harms

- Test-related discomfort associated with the cytological sampling procedure and follow-up tests and procedures for cytological abnormalities such as colposcopy and biopsy, many of which will be found to be normal.
- Harms of cervical procedures for treatment of pre-invasive cancer, some of which pre-cancers will spontaneously regress, with the risk of adverse pregnancy outcomes including perinatal mortality.

Psychological harms

- Test-related anxiety and impact on relationships following an abnormal cytology result
- Patient inconvenience of screening and follow-up procedures.

Summary of harms associated with treatment of pre-invasive cancer

This review does not systematically assess the harms associated with the treatment of pre-invasive cancer, and therefore a summary of findings from the most recent systematic review to assess these harms is provided.

The Peirson review for the Canadian Task Force on Preventive Health Care (CTFPHC) (Peirson et al. 2012) addressed the following question: What are the harms of treatment of cervical cancer?

- The following harms were considered:
    - harms of colposcopy
    - harms of biopsy: cone biopsy (immediate and late; pre-term labour, miscarriage) and loop electrosurgical excision procedure (LEEP)/ large loop excision of the transformation zone (LLETZ) (immediate and late effects)
    - harms of treatment of cervical cancer: total hysterectomy (incontinence, infection, hospitalisation) and radiotherapy.

The literature search covered the period from 2005 to February 2011 and included studies of any design.

Peirson (2012) identified one study reporting on the physical effects of colposcopy alone, colposcopy and cervical punch biopsies, and colposcopy and LLETZ: the TOMBOLA trial (TOMBOLA 2009). Of the 401 women who received a colposcopy, only 14–18% reported pain, bleeding or discharge, all ranging in severity from mild to moderate (no severe after-effects were reported). Twenty-nine per cent of women reported a change in their first post-colposcopy menstrual period. Women in the LLETZ group (N=185) experienced a higher frequency of all physical after-effects than women in the biopsy (N=165) group or colposcopy alone group.

Two systematic reviews with meta-analyses (Arbyn et al. 2008; Kyrgiou et al. 2006) suggest that cold-knife conisation increased the risk of:

- perinatal mortality (7 studies: RR 2.87, 95% CI 1.42–5.81)
- severe pre-term (<32–34 weeks) delivery (five studies: RR 2.78, 95% CI 1.72–4.51)
- extreme pre-term (<28–30 weeks) delivery (four studies: RR 5.33, 95% CI 1.63–17.40)

- severe (<2000g) low birth weight (1 study: RR 2.86, 95% CI 1.37–5.97)
- caesarean section (four studies: RR 3.17, 95% CI 1.07–9.40).

One study found that laser conisation increased the risk of severe (<2000g) low birth weight (RR 3.50, 95% CI 1.06–11.53) and extreme (<1500g) low birth weight (RR 10.00, 95% CI 1.19–83.84). In meta-analysis the risk of preterm delivery was non-significantly increased (six studies: RR 1.71, 95% CI 0.93–3.14).

The effects of LEEP/LLETZ on pregnancy outcomes were assessed by nine studies, including two meta-analyses (Arbyn et al. 2008; Kyrgiou et al. 2006) and seven primary studies (Samson et al. 2005; Ortoft et al. 2010; Noehr et al. 2007; Noehr et al. 2009; Tan et al. 2004; Paraskevaidis et al. 2002; Sadler et al. 2004).

The most recent meta-analysis (2008) found no significant increase in risk of perinatal mortality, pre-term (<32–34 weeks) delivery or severe low birth weight (<2000g). Four of the seven primary studies found an increased risk of pre-term (<37 weeks) delivery among the LEEP/LLETZ patients, as did the earlier (2006) meta-analysis (eight studies: RR 1.70, 95% CI 1.24–2.35).

The 2008 meta-analysis calculated the absolute risks to derive the number needed to treat to observe obstetric harm in one treated woman (Number Needed to Harm (NNTH)). For the outcomes of perinatal mortality this number was:

- 71 (six studies) for knife conisation
- 67 (three studies) for laser conisation
- 500 (seven studies) for LLETZ.

For the outcome of severe preterm delivery (<32–34 weeks) the NNTH was:

- 30 (five studies) for knife conisation
- 167 (one study) for laser conisation
- 143 (four studies) for LLETZ.

For the outcome of severe low birthweight (<2000g) the NNTH was:

- 16 (1 study) knife conisation
- 14 (1 study) laser conisation
- 106 (1 study) LLETZ.

Overall, the harms associated with LLETZ were much less common than those resulting from cold knife or laser conisation.

Two studies addressing the outcome of sexual functioning for LEEP patients were also identified (Hellsten et al. 2008; Inna et al. 2010). One study found no differences between mean scores of LEEP and non-LEEP patients with respect to psychosexual functioning (including frequency of intercourse, spontaneous interest, sexual arousal, orgasm, lubrication, dyspareunia, negative feelings). One study found no differences in terms of frequency of sexual intercourse, dysmenorrhea, dyspareunia, or post-coital bleeding; however, significant differences between groups were evident in terms of overall satisfaction ($p=0.01$), vaginal elasticity ($p=0.03$) and orgasmic satisfaction ($p=0.03$).

One study (Sharp et al. 2011) compared LLETZ to biopsy and recall for the outcomes of anxiety and depression. No differences between groups were evident with respect to prevalence of significant depression. Overall, there were also no differences between groups in anxiety (except prior to the procedure—anxiety was higher pre-LLETZ than pre-biopsy).

No evidence was identified pertaining to the harms of total hysterectomy as a treatment for cancer, and the authors concluded that the evidence pertaining to the harms of radiotherapy as treatment for cancer was unclear.

Patient values and preferences around screening

This section includes the data from the Peirson (2012) review and two systematic reviews that have been recently published on women's preferences in the management of low-grade abnormal findings.

The Canadian review (Peirson et al. 2012) examined the following patient values and preferences in the context of cervical screening: preferences regarding screening intervals; preferences regarding healthcare providers; cultural views affecting participation; and factors related to screening/intention to be screened.

Screening interval

Two studies addressed patient preferences regarding the screening intervals. A US study suggests that 78.8% of women believed that a woman should obtain a Pap test annually (although 63.1% were willing to attend screening every three years) (Meissner et al. 2010). A survey of 167 Australian women aged 18–69 showed more willingness to attend annual rather than biennial screening, and less willingness to attend at three- or five-year intervals. This however was driven by GPs adhering to national guidelines, as the study concluded that a change to the recommended screening interval would be followed by both women and general practitioners, and be unlikely to have an impact on screening attendance (Fiebig et al. 2009)

Similar findings were reported by Wordsworth (2006) who undertook a study to ascertain women's preferences on different screening intervals. This study was not included in the Peirson review (2012) but was identified during the literature search phases of this project. It was noted that changes to the frequency of screening were unlikely to affect attendance, with the exception of women over the age of 50 years who had a stronger preference for more frequent screening than younger women.

In interpreting this evidence it is important to note that none of the three studies asked explicit questions regarding the trade-off between frequency of screening and reduction in cancer risk. In addition, the issue of HPV testing was not addressed.

Preferences regarding healthcare providers

According to six studies (including one systematic review) carried out across groups of women of various ethnicities (including Jordanian, Crow Indian, African Caribbean, Pakistani, Greek, Arabic, and Australian women), patients prefer having tests performed by female professionals (physicians, gynaecologists, etc.) (Fiebig et al. 2009; Ackerson et al. 2008; Barghouti et al. 2008; Smith et al. 2008; Johnson et al. 2008; Thomas et al. 2005). Male providers were found to increase the embarrassment of women (one study) and make them less likely to screen (one study).

The cultural views affecting women's participation in screening were examined by six studies (Thomas et al. 2005; O'Brien et al. 2009; Chang et al. 2010; Brotto et al. 2008: Ji et al. 2010; Wang et al. 2008). The following cultural views affected women's participation in screening: embarrassment

associated with testing; fear of cancer; inadequate information about cancer; preference for female health providers (due to cultural beliefs); and desire to adhere to cultural norms around preservation of virginity. Four studies identified acculturation as a barrier to cervical screening (a greater degree of acculturation to Canadian and American culture translated into greater likelihood of screening participation and reproductive health knowledge).

Participation

A number of factors affected the participation in screening or the intention to participate in screening. Peirson et al. (2012) identified the following recurring factors (across 27 studies)[1]: time as a barrier to participation (including other commitments; inconvenience of appointments; 'not getting around to it'), history of trauma/abuse, previous negative healthcare or Pap experiences, fear of the results/ preference for not knowing, fear of not detecting the disease, and misperception of risks/benefits of screening (eg regarding risks associated with age, sexual activity, family history, general good health). Salient socio-cultural factors included: fatalistic attitudes; embarrassment; fear of pain; anxiety/ stress related to diagnosis; lack of provider recommendation; presence of male provider; distrust in the healthcare system; difficulty finding interpreters; lack of knowledge; and belief that screening is unnecessary in the absence of illness.

Management of low-grade abnormal findings (including HPV triage)

Two systematic reviews have recently been published on women's preferences in the management of low-grade abnormal findings (Frederiksen et al. 2012; Sadique & Legood 2012). Both systematic reviews include Australian data (Howard et al. 2008; McCaffery et al 2008).

The review by Frederiksen (2012) included 13 studies. To be included, studies had to ask women about their preferences between active follow-up and observation in the management of low-grade abnormalities. Active follow-up was defined as immediate colposcopy, repeated testing in less than 6–12 months and HPV triage. Observation was considered repeated Pap tests or HPV testing and less frequent colposcopy. In eight of the thirteen studies women had low-grade abnormalities, whereas in five studies women's preferences were based on a hypothetical situation. Out of the eight studies of women with abnormal findings, five asked women about their preferences prior to any management or treatment. It is these five studies that the authors state are the most applicable in terms of determining women's preferences. In all five studies it was reported that the majority of women preferred active follow-up (mostly immediate colposcopy) to observation (predominately as repeated Pap tests). However, women appeared to be more willing to accept observation if reassured of the low risk of cervical cancer. One of the difficulties in interpreting this data is the heterogeneity between the studies. There was substantial variation in terms of the women included and the methods used to elicit preferences.

Sadique and Legood (2012) also reviewed the literature on women's preferences in the management of low-grade abnormalities. This review included nine studies, six of which were also included in the review by Frederiksen (2012). The authors note in this review that the findings of the studies differed depending on the methodological approach. However, they do state that when preferences were elicited from women undergoing management of low-grade abnormalities (in contrast to

[1] These included: two systematic reviews (Fiebig 2009; Johnson 2008), one integrated review (Ackerson 2007), and 24 primary studies (Kulasingam 2009; Ackerson 2008; Barghouti 2008; Thomas 2005; Ji 2010; Xiong 2010; Amankawah 2009; Waller 2009; Wong 2008; Wake 2009; Denberg 2005; Blomberg 2008; Farley 2002; Guilfoyle 2007; Kuitto 2010; Carter 2002; Lopez 2006; Nelson 2002; Hoyo 2005; Behbakht 2004; Eiser 2002; Oscarsson 2008; Jepson 2007; Vanslyke 2008).

hypothetical scenarios) women were more likely to opt in favour of active treatment (HPV testing and colposcopy) rather than repeat cytology.

HPV triage (Australian data)

As mentioned previously, both reviews (Frederiksen et al. 2012; Sadique & Legood 2012) included the Australian studies of Howard et al. (2008) and McCaffery et al. (2008) in their analysis of women's preferences. McCaffery et al. have also published a more recent paper (2010) that was excluded by Frederiksen et al. (2012) as it was considered to contain duplicate information to that in the earlier publication. Briefly, the trial conducted by McCaffery et al. (2010) was a pragmatic RCT that compared psychosocial outcomes in women 16–70 years with a borderline (ASCUS) cervical smear randomised to either HPV testing, repeat smear, or either test based on informed choice. It was reported that while the psychosocial effect for women allocated to HPV triage was worse, this reduced over the study period, and by 12 months the outcomes appeared no better than those who underwent repeat cytology testing.

Methodology

Objective

The main objective of the review is to conduct a systematic review and comparative analysis of cervical screening technologies, cervical screening age range and intervals, and identify potential cervical screening pathways.

Clinical decision pathway

The algorithm for the current cervical screening program is presented in Figure 6.

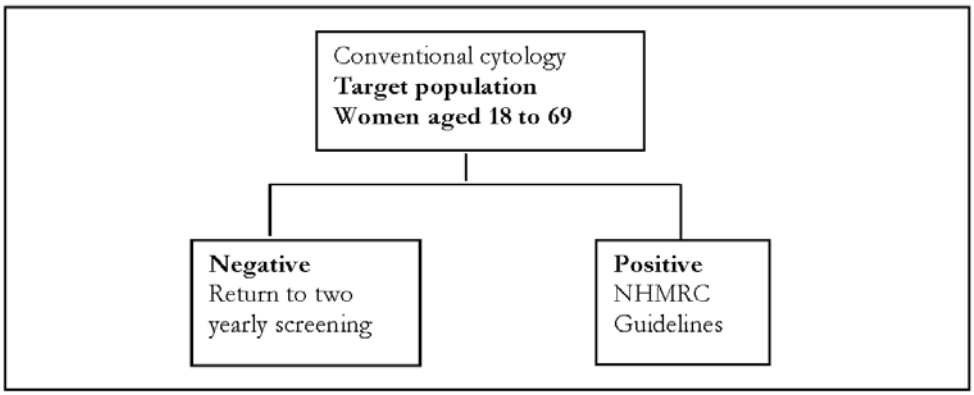


Figure 6 Algorithm for the current cervical screening program

Comparator

The comparator for any alternative cervical screening pathway (primary questions) is the current pathway promoted by the NCSP. All women between the ages of 18 (or one to two years after first having sexual intercourse, whichever is later) and 70 years of age should be screened every two years using conventional cytology.

The Renewal Steering Committee (following finalisation of the DAP), advised that the comparator for secondary questions are the screening strategies considered in the relevant primary question.

The reference standard

For studies to provide adequate information about test accuracy, a reference standard should be applied as a minimum to positive or discordant tests. Colposcopy with biopsy (threshold for positive histology CIN2+/CIN3+) was considered the most valid reference standard to determine the true disease status of patients with a positive test (pLSIL, dLSIL, pHSIL, HSIL or squamous cell carcinoma (SCC)). Clinical follow-up with repeat cytology at one year was considered the most valid reference standard for normal test results. Adjudicated cytology was considered a suboptimal reference standard for patients with either test result.

For estimates of absolute sensitivity and specificity, verification of negative test results is required. Where this has not been performed, relative measures of test accuracy (relative sensitivity and specificity) are required. In paired studies, with confirmation only of discordant test results, where

neither absolute or relative sensitivity and specificity can be determined, the ratio of additional true positive (TP) to false positive (FP) findings, or the incremental rate of TPs can often be determined to allow a comparison of test performance.

Methodological issues

The key methodological issues relevant to the assessment of evidence for cervical screening programs have been outlined by IARC (IARC 2005) and further described by Arbyn and colleagues for the European Guidelines for Quality Assurance in Cervical Cancer Screening (Arbyn et al. 2008, 2009). As outlined in these documents, the effectiveness of cervical screening to reduce the incidence and mortality of cervical cancer depends on a set of factors:

- the sensitivity of the primary screening test
- the clinical thresholds used for follow-up and treatment
- the sensitivity of triage and diagnostic work-up for abnormalities detected
- participation rates
- participant adherence to follow-up
- the age group targeted
- the screening interval
- the natural history of the disease; where essential elements are the rates of onset of precursor lesions; the progression and regression rates of these precursor lesions; and the distribution of their sojourn times (Arbyn et al. 2009).

A methodological framework has recently been developed to describe the evidence requirements for assessment of the impact of a screening strategy on patient and population relevant outcomes, including consideration of the harm and benefit trade-off and cost-effectiveness (Arbyn et al. 2013).

A major challenge for reviewers is the need to systematically identify, appraise and synthesise evidence that is continually updated with ongoing advances in cervical screening technologies. To help address this challenge, an international collaborative research team of clinical trialists and experts in systematic reviews has been established to conduct future meta-analyses of the HPV screening trials and to synthesise evidence on other new cervical cancer prevention strategies (Arbyn et al. 2009).

Mathematical models are used to incorporate the international evidence, and apply it to local settings using relevant epidemiological data on cervical cancer and pre-cursors and screening program participation to estimate outcomes for specific strategies.

Assessing the performance of different screening program strategies for age range and screening interval

Despite the strong body of evidence demonstrating the effectiveness of cytology-based screening to reduce the incidence and mortality of cervical cancer (IARC 2005), evidence to determine the optimal age range and interval for screening is limited and screening programs in different countries use a range of different strategies (Appendix C).

The optimal screening age range and interval will represent the strategy that provides the most favourable trade-off between the benefits of detecting more pre-cancers in order to reduce cancer incidence and mortality; versus the harms of performing tests and procedures for cytological abnormalities that are subsequently found to be normal and the harms of treatment for

precancerous changes, some of which may spontaneously resolve. A consideration of this trade-off must also take resource use and costs into account. The trade-off between benefits and harms may be unfavourable for subpopulations with a lower baseline risk of disease events, a higher risk of harms, or because the relative effectiveness of screening is lower in the subpopulation.

Ideally, RCTs of screening strategies that have used different screening intervals and report subgroup analysis for different age groups with long-term follow-up providing data on cervical cancer incidence, mortality and test/treatment-related harms would be available to provide 'direct' evidence about the optimal screening strategy. In the absence of such trials, standard epidemiological principles for clinical research can be used to define and assess relevant evidence to estimate the net clinical benefit of different screening strategies.

Using these principles to examine the optimal age range for screening, where the relative effectiveness of screening has been established in a broad age group of women, but less evidence is available for the two age groups of specific interest for this review (start age 20 versus 25 years; exit age 65 versus 70 years), conclusions can also be informed by evidence about age-specific differences in the risk and natural history of cervical cancer, population characteristics, the harms of screening procedures and program participation. Each of these factors may limit the generalisability of estimates of screening effectiveness from the broad screening population to women aged less than 25 years and those over 65 years.

To examine the optimal screening interval, evidence is needed to assess the relative effectiveness of screening versus no screening when using different screening intervals; with subgroup analyses to determine whether this difference varies for different age groups.

In the absence of direct evidence, studies that report the risk of cervical cancer following a negative screening test at each of the different intervals being considered can be used to determine the longest interval that is still associated with a desired reduction in the risk of cervical cancer compared to no screening.

Studies that report age-specific risks are needed to consider the use of different screening intervals for different age groups. The screening interval chosen will also depend on whether cytology or HPV is used as the primary screening test because the risk of developing a pre-invasive and invasive cancer following a negative cytology test is different to a negative HPV test (Kitchener et al. 2011b).

The trade-offs between more versus less intensive screening intervals can be examined in a modelled cost-effectiveness analysis for different screening intervals.

Comparing the performance of different screening tests

Where comparative accuracy of two alternative replacement test strategies is required, studies can either be performed with both tests conducted in all patients, or by randomising patients to either test (Bossuyt et al. 2006). Conducting both tests in all patients (a paired design) has advantages as the patient population undergoing each test is identical. In comparing the accuracy of conventional Pap tests and LBC, the 'split-sample' technique is frequently used to perform both tests in the same patients. This approach is believed to disadvantage the technical performance of LBC as more cells are applied to the smear sample than rinsed into the LBC transport medium vial. Therefore, for the current review, randomised studies are believed to provide a more valid estimate of comparative accuracy than paired studies.

The sensitivity of a screening program differs to that of the one-time use sensitivity of the screening test. Screening program sensitivity is dependent on the sensitivity of the test with

repeated use over multiple rounds, as well as screening program participation rates, accuracy of follow-up procedures, and compliance with follow-up and treatment.

Classification and management of cytological and histological abnormalities

Different systems are used for classifying cytological and histological abnormalities in cervical screening. In Australia, cytological abnormalities are classified using the Australian Modified Bethesda System (AMBS). Under this system, cytological abnormalities of squamous cells are classified as high-grade squamous intraepithelial lesions (HSILs), possible HSILs (pHSILs), low-grade squamous intraepithelial lesions (LSILs) or possible LSILs (pLSILs) (Table 7). The international literature most commonly uses the US Bethesda System, which uses a slightly different terminology. pLSIL is equivalent to atypical squamous cells of undetermined significance (ASCUS) under the US Bethesda System (Table 7).

In Australia, women with HSIL are referred to a specialist for examination of the cervix by use of a colposcope, in a procedure called a 'colposcopy'. Abnormal lesions identified at colposcopy are biopsied and classified as CIN grades 1–3 on the basis of the histological findings. Although it was originally believed that neoplastic cellular changes occurred along a continuum from CIN1 to 2 to 3, CIN1 is now regarded as a manifestation of the HPV infective process, rather than as the first step in the neoplastic process.

HPV infection of the cervix is usually asymptomatic, and most infections are transient. HPV infection may not cause any change in cell morphology or it may cause the cytopathic effect previously recognised as mild dysplasia and classified as CIN1. Thus, CIN1 lesions are now monitored by repeat cytology with the expectation that the cellular changes will regress when HPV infection resolves. In a small proportion of women, persistent HPV infection may occur.

Persistent infection with oncogenic genotypes precedes precancerous changes, which are classified as CIN2 (moderate dysplasia) or CIN3 (severe dysplasia). These lesions are treated by ablative therapy to prevent progression to invasive cancer. It is now accepted that CIN2 or 3 can occur de novo, rather than as a continuum from CIN1 lesions. A trial-based quality control assessment of community pathology biopsy diagnoses has demonstrated that the detection of CIN2 has poor reproducibility compared to the detection of CIN3, with 56% of 523 CIN2 cases reclassified as CIN3 (27%) or <CIN2 (29%) at the quality control assessment (Castle et al. 2007). The authors suggested that this evidence indicates that CIN2 represents a mix of HPV infection and CIN3, and that CIN3 is the true precursor to cancer (Castle et al. 2007).

Women with a cytological finding of possible or definite LSIL are managed more conservatively, with cervical cytology repeated at 12 and 24 months and referral for colposcopy only if these lesions are persistent, because the majority represent an infective process due to HPV and will resolve spontaneously without treatment. However, around 20% of LSIL cases will be confirmed as CIN2–3 at histology if immediate colposcopy and biopsy are performed (pooled prevalence from 10 studies: 18.8% [95% CI 1.24–25.2]; Arbyn et al. 2006).

Table 7 Comparison of the Australian Modified Bethesda System, 2004, and the US Bethesda System, 2001

Australian Modified Bethesda System	US Bethesda System
Squamous abnormalities	***Squamous abnormalities***
Possible low-grade squamous intraepithelial lesion (pLSIL) Low-grade squamous intraepithelial lesion (LSIL) Possible high-grade squamous intraepithelial lesion (pHSIL) High-grade squamous intraepithelial lesion (HSIL) Squamous cell carcinoma (SCC)	Atypical squamous cells, undetermined significance (ASCUS) Low-grade squamous intraepithelial lesion Atypical squamous cells, possible high-grade lesion (ASC-H) High-grade squamous intraepithelial lesion Squamous cell carcinoma
Glandular abnormalities	***Glandular abnormalities***
Atypical endocervical cells of undetermined significance Atypical glandular cells of undetermined significance (AGUS) Possible high-grade glandular lesion Endocervical adenocarcinoma in situ (AIS) Adenocarcinoma	Atypical endocervical cells, undetermined significance Atypical glandular cells of undetermined significance Atypical endocervical cells, possibly neoplastic Endocervical adenocarcinoma in situ Adenocarcinoma

Source: Extracted from the NHMRC *Screening to prevent cervical cancer guidelines* (NHMRC 2005).

Natural history of cervical cancer

Cervical cancer arises via a series of four steps: HPV infection; viral persistence; progression of a clone of persistently infected cells to pre-cancer; and invasion. Backward steps also occur, namely clearance of HPV infection and regression of pre-cancer to normality (Schiffman et al. 2007). The prevalence of transient HPV infections is highest among young women in the first few years after the initiation of sexual activity. As genital HPV is primarily sexually acquired and readily transmissible, precancerous cervical lesions (CIN2-3) peak a decade later. In a small proportion of women, precancerous cervical lesions progress to invasive cancer over a period of 10 to 20 years, peaking among women aged 40 to 50 years.

Pre-cancer is the target of screening and preventative treatment but is heterogeneous. CIN2 lesions can be caused by HPV types rarely found in cancer and have a sizeable regression potential whereas CIN3 lesions share the same HPV type spectrum and causal cofactors as cancer, although those detected by screening may be small and not certain to progress. Carcinoma in situ is the most certain surrogate for invasive cervical cancer (Schiffman et al. 2007). The use of precancerous surrogates in test performance studies and trials is a limitation of the evidence and the degree of benefit in preventing invasive cervical cancer is usually not known.

Most cervical HPV infections are cleared or suppressed within one to two years of exposure and the median time to clearance of HPV infections detected during screening studies is 6–18 months (Schiffman et al. 2007). Rates of progression and regression of pre-cancer were summarised in the Agency for Healthcare Research and Quality (AHRQ) review (Vesco et al. 2011) which reported progression of CIN3 to cancer at 31% in 30 years (from an unethical New Zealand study, subsequently reported by McCredie et al. 2008), and other estimates of progression of 20–30% over 5–10 years (Chang 1990; Kinlen et al. 1978). They also cited evidence about CIN regression, including a review of studies published up to 1990 that reported on regression, persistence and progression over follow-up from 1–25 years for CIN1 (57%, 32%, 1% respectively); CIN2 (43%, 35%, 5% respectively) and CIN3 (32%, 56% and >12% respectively) (Ostor 1993). A large Canadian cohort study that managed CIN2 conservatively reported 0.3% progression to invasive cancer within two years, 0.7% within five years and 1.2% within 10 years; with regression to normal cytology reported at 6.9% within two years, 29.0% within five years and 53.7% within 10

years (Holowaty et al. 1999). Rates of CIN3 progression to invasive cancer were higher (1.6% within two years, 2.6% within five years and 9.9% within 10 years (Holowaty et al. 1999)).

Because cervical cancer has a long preclinical period with regression as well as progression, and because studies rely on surrogate precancerous outcomes, there is a risk that outcomes which appear to be beneficial may actually represent the diagnosis and treatment of disease which would otherwise regress; or the earlier diagnosis of disease, with unknown benefit, which would have been detected in a subsequent screening round anyway.

Data from the NCSP show that abnormalities are most common in younger women, due to HPV infections that occur frequently after sexual debut. Abnormal cytology is highest in women under 20 and in those aged 20–24, while high-grade abnormalities are relatively low in women under 20 and peak in women aged 20–29 (Table 8). Detection rates of both low-grade and high-grade abnormalities then decrease with increasing age, only increasing slightly in women aged 70 or over (Table 8; Table 9; AIHW 2013). The number of histology tests per 100 cytology tests peaks in the 20–29 year age group at 5.2 and then steadily declines to 1.9 at 60–64 (AIHW 2013). Cancer incidence and mortality is low in women aged less than 30 years (see Age-stratified incidence and mortality, page 10).

Table 8 Number and proportion of abnormal cytology and high-grade abnormalities detected by histology, by age, 2010

Age group (years)	Screening outcomes					
	Number of women screened	Number abnormal cytology	Proportion abnormal cytology per 100 women screened	Number high-grade histology	Proportion high-grade histology per 1000 women screened	Number carcinoma of the cervix
<20	53,297	7,448	14.0	551	10.3	0
20–24	181,124	25,721	14.2	4,879	26.9	26
25–29	227,113	23,357	10.3	6,263	27.6	89
30–34	234,547	16,341	7.0	4,410	18.8	84
35–39	254,032	12,676	5.0	2,841	11.2	139
40–44	236,562	9,419	4.0	1,519	6.4	84
45–49	229,422	7,923	3.5	1,032	4.5	96
50–54	199,794	5,107	2.6	513	2.6	84
55–59	163,697	3,338	2.0	342	2.1	86
60–64	135,541	2,150	1.6	190	1.4	70
65–69	84,810	1,229	1.4	115	1.4	60
70–74	20,035	445	2.2	46	2.3	52
75–79	4,621	140	3.0	31	6.7	25
80–85	1,696	73	4.3	12	7.1	41
85+	679	42	6.2	7	10.3	22
All ages	2,026,980	115,409	5.7	22,757	11.2	967
Ages 20–69	1,946,642	107,261	5.5	22,104	11.4	818

Data are from all jurisdictions. Abnormal cytology and high-grade histology include both squamous and endocervical abnormalities. High-grade abnormalities are defined as HS03 (CIN NOS, CIN2 and CIN3) and HE03 (endocervical dysplasia and adenocarcinoma in situ). Carcinoma of the cervix includes squamous cell carcinoma, adenocarcinoma, adenosquamous carcinoma, and other carcinoma of the cervix.

Source: AIHW (2013). *Cervical screening in Australia 2010–2011*. Cancer series 76. Cat. no. CAN 72. Canberra: AIHW data (additional material).

Table 9 Number of CIN2 and CIN3 detected by histology and as a proportion of all histology tests, by age, 2010

Age group (years)	Screening outcomes				
	Number of women screened	Number CIN2	Proportion CIN2	Number CIN3	Proportion CIN3
<20	24,679	146	20.9	102	14.6
20–24	90,671	1,123	21.5	1,036	19.8
25–29	112,022	1,288	21.0	1,495	24.3
30–34	113,718	804	17.2	1,055	22.6
35–39	123,427	469	11.2	690	16.4
40–44	117,374	272	6.5	347	8.3
45–49	112,377	198	4.5	232	5.3
50–54	98,344	88	3.1	104	3.6
55–59	81,210	58	3.2	86	4.8
60–64	66,898	21	1.6	50	3.9
65–69	41,843	17	1.9	32	3.5
70–74	9,643	9	1.9	16	3.4
75–79	2,188	3	1.1	11	4.2
80–85	823	0	0.0	7	4.5
85+	350	1	1.5	1	1.5
All ages	995,568	4,497	12.3	5,264	12.2
Age 20–69	957,884	4,338	10.1	5,127	12.4

Data are only included from jurisdictions who can distinguish between CIN2 and CIN3.

Source: AIHW 2013. *Cervical screening in Australia 2010–2011*. Cancer series 76. Cat. no. CAN 72. Canberra: AIHW data (additional material).

Research questions

The research questions as defined in the Decision Analytic Protocol (DAP) are listed below and a comparison across the three cervical screening scenarios and the current program is presented in Table 10.

Primary question 1

What is the comparative safety, effectiveness and cost-effectiveness of conventional cytology, using the IARC recommendations for age range and interval, compared with the protocol used in the current Australian cervical screening program?

Secondary questions:

1. What is the safety, effectiveness and cost-effectiveness of using one HPV test for women exiting the program at age 65 years and over, compared with the existing protocol?

2. What is the cost-effectiveness of the pathway if an invitation/recall system was introduced compared with the existing overdue reminder system without invitation? (Addressed in the modelling section only, accompanying report 'Effectiveness modelling and economic evaluation in the Australian setting')

Primary question 2

What is the comparative safety, effectiveness and cost-effectiveness of either filtration or cell enrichment LBC (using the IARC recommendations for age range and interval for cytology), compared with the protocol used in the current Australian cervical screening program?

Secondary questions:

1. What is the comparative safety, effectiveness and cost-effectiveness of using automated image analysis under Question 2 parameters, compared with the existing protocol?

2. What is the comparative safety, effectiveness and cost-effectiveness of adding an HPV test to triage women with pLSIL/LSIL, compared with the existing protocol?

3. What is the comparative safety, effectiveness and cost-effectiveness of using one HPV test for women exiting the program at age 65 years and over, compared with the existing protocol?

4. What is the comparative safety, effectiveness and cost-effectiveness of undertaking a colposcopy immediately in comparison to delaying the test in women who have pLSIL/LSIL cytology and a positive HPV test?

Primary question 3

What is the comparative safety, effectiveness and cost-effectiveness of HPV testing as the primary screening test in women aged 25 to 65 years every five years, compared with the protocol used in the current Australian cervical screening program?

Secondary questions:

1. What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a co-test in the above scenario, compared with the existing protocol?

2. What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a reflex test to triage women with positive HPV test results, compared with the existing protocol?

3. What is the comparative safety, effectiveness and cost-effectiveness of undertaking a colposcopy immediately in comparison to delaying the test in women who have pLSIL/LSIL cytology and a positive HPV test?

4. What is the comparative safety, effectiveness and cost-effectiveness of including self-collected samples for HPV testing for underscreened and unscreened women to supplement the organised screening program using practitioner-collected HPV samples, compared with the existing protocol?

5. What is the comparative safety, effectiveness and cost-effectiveness of referring women positive for HPV 16/18 +/–45 using partial genotyping systems at primary screening, immediately to colposcopy and performing cytology triage on women positive for other oncogenic types?

Following finalisation of the DAP, the Renewal Steering Committee advised that the comparator for secondary questions are the screening strategies considered in the relevant primary question.

Table 10 Summary of cervical screening scenarios to be compared with the current cervical screening program

		Comparator Current program	Scenario 1	Scenario 2	Scenario 3
Primary question	Primary screening test	Conventional cytology	Conventional cytology	LBC (cell filtration and cell enrichment separately)	HPV testing (including information on genotyping separately)
	Age range	Women aged 18–69 years	Women aged 25–64 years (IARC recommendations)	Women aged 25–64 years (IARC recommendations)	Women aged 25–64 years (IARC recommendations)
	Interval	2 years	3 years (aged 25–49) and 5 years (aged 50–65) (IARC recommendations)	3 years (aged 25–49) and 5 years (aged 50–65) (IARC recommendations)	No less than 5 years (a range of intervals should be considered)
Secondary questions	Triage options**	pLSIL/LSIL result N/A (as per NHMRC Guidelines)	pLSIL/LSIL result N/A (as per NHMRC Guidelines)	pLSIL/LSIL result +/– reflex HPV testing	HPV positive result 1) +/– LBC co-testing 2) +/– LBC reflex testing
	Additional technology	N/A	N/A	+/- Automated image analysis	+/- Automated image analysis
	Exit strategy	Must have two normal cytology tests within the last 5 years	HPV test at age 64 years	HPV test at age 64 years	HPV test at age 64 years
	Self-collection	N/A	N/A	N/A	YES
	Invitation and recall system	N/A (overdue reminders only)	YES	YES	YES

** In this assessment triage tests include additional tests undertaken in the laboratory, using the original sample, which will assist in making a final recommendation on the index test, based on the combined results. This does not involve tests for the follow-up of an abnormal result.

Review of literature

Existing health technology assessment (HTA) reports and systematic reviews

Prior to conducting the specific literature searches for the questions outlined in scenarios 1–3, a review of existing HTAs and systematic reviews was undertaken. Searches were carried out on 31 October 2012. Websites of international health technology assessment agencies were searched, as well as the Cochrane Database of Systematic Reviews, and the NHS Centre for Reviews and Dissemination database (including DARE, NHS EED and HTA). The following search terms were used: LBC, HPV, "cervical screening." Search restrictions were: LBC (2008-present), HPV (no date restrictions), "cervical screening" (2008–present). The 2008 date was chosen as the cut-off for LBC and cervical screening searches as the last MSAC reviews of those areas were conducted in 2009, with searches to 2008. Twenty-two relevant HTAs, systematic reviews and/or meta-analyses were identified; they are listed in Appendix G.

The most recent high-quality systematic reviews identified in that search which addressed the primary research questions were the AHRQ HTA conducted for the US Preventive Services Task Force (USPSTF; Vesco et al. 2011); a Canadian systematic review conducted to inform revisions to the Canadian Services Task Force Recommendations (Peirson et al. 2012); a systematic review conducted to inform the American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer (Saslow et al. 2012); and a meta-analysis of HPV screening strategies by Arbyn et al. (2012).

The AHRQ report was considered the most comprehensive. It reviewed evidence to September 2010 and conducted an ancillary search which identified additional studies published to August 2011. This report considered evidence which addressed many of the research questions of the current review including:

- age of initiation of screening
- manually read LBC compared to conventional cytology
- HPV testing for primary screening, with and without LBC triage of positive results
- harms of LBC
- harms of HPV testing.

As such this report was used as the source of primary literature to 2010–2011 for questions included in the AHRQ report and also included in this review. A separate literature search was not undertaken to identify studies that addressed the secondary research questions. Studies identified in the search for studies addressing the primary research questions were screened to identify studies that specifically addressed these questions and that could inform the model.

Literature sources and search strategies

Conventional cytology – age range and screening interval (scenario 1)

MSAC has not previously reviewed the age range and interval for cervical screening.

Primary research question

To design the literature search, the primary research question for scenario 1 was broken down into four sub-questions:

1.1 What is the comparative safety, effectiveness and cost-effectiveness of conventional cytology using a start age at 25 years versus the current start age of 18–20 years for sexually active women?

1.2 What is the comparative safety, effectiveness and cost-effectiveness of conventional cytology using an exit age of 65 years versus the current exit age of 70 years for women who have had two negative tests within the last five years?

1.3 What is the comparative safety, effectiveness and cost-effectiveness of conventional cytology using a three-year screening interval for women aged 25–50 years versus the current 2-yearly screening interval for women of all ages?

1.4 What is the comparative safety, effectiveness and cost-effectiveness of conventional cytology using a five-year interval for women aged 50–65 years versus the current 2-yearly screening interval for women of all ages?

Sources used to identify eligible studies included:

- The IARC handbook of cancer prevention: Cervix cancer screening, 2005.
- A search of the medical literature to identify (i) systematic reviews published in the period 2004 to November 2012 that assessed screening age range or screening interval; and (ii) primary studies published in the period 2010 to November 2012.

The 2004 start date for the search for HTAs and systematic reviews was chosen to update the evidence available to the IARC working group meeting in 2004, prior to publication of the IARC recommendations in 2005. The 2010 start date for the primary studies search was chosen to update the last search date (February–August 2011) of three recent eligible systematic reviews identified (Vesco et al. 2011; Peirson et al. 2012; Saslow et al. 2012). Reference lists and citations of included studies were also checked.

Secondary research questions

Mathematical modelling was be required to assess secondary research questions for scenario 1 on the use of i) an HPV test as an exit strategy for women over 65 years; and ii) an invitation/recall system versus the existing overdue reminder system. Model inputs were based on existing evidence about the accuracy of HPV screening; the incidence of HPV and the risk of cervical cancer following a negative HPV test in older populations; and published data about the impact of invitation/recall systems. Details of the modelling inputs and assumptions are provided in the accompanying economic evaluation of this assessment ('Effectiveness modelling and economic evaluation in the Australian setting').

Search strategies

Searches for HTAs and systematic reviews are described on page 38. Searches for additional primary studies were conducted in EMBASE, Medline and PubMed as listed in Table 11.

Table 11 Electronic databases searched for primary studies (scenario 1)

Database	Period covered
EMBASE	2010–2012 (search performed on Nov 20, 2012)
Medline (via Embase.com)	2010–2012 (search performed on Nov 20, 2012)
PubMed	2010–2012 (search performed on Dec 6, 2012)

The search terms used to identify primary studies were adapted from the AHRQ (Vesco et al. 2011) search strategy and are listed in Appendix F.

Liquid-based cytology (scenario 2)

The medical literature was searched to identify relevant studies and reviews for the period between January 2008 and November 2012.

A search for manually read LBC primary studies was conducted from January 2010 to November 2012 to update the most recent high-quality systematic review (Vesco et al. 2011) including studies on the comparative effectiveness of LBC and conventional cytology. A search to identify primary studies of automated image analysis was conducted from January 2008 to November 2012 to update the 2009 MSAC review.

Searches were conducted in EMBASE, Medline and PubMed as listed in Table 12.

Table 12 Electronic databases searched for primary studies (scenario 2)

Database	Period covered (LBC)	Period covering (automated)
EMBASE	2010–2012 (search performed on 21 November 2012)	2008–2012 (search performed on 21 November 2012)
Medline (via Embase.com)	2010–2012 (search performed on 21 November 2012)	2008–2012 (search performed on 21 November 2012)
PubMed	2010–2012 (search performed on 6 December 2012)	2008–2012 (search performed on 6 December 2012)

The search terms used to identify primary studies were adapted from the AHRQ (Vesco et al. 2011) search strategy and are listed in Appendix F.

HPV testing (scenario 3)

The medical literature was searched to identify relevant studies and reviews for the period between 2010 and December 2012. Searches for additional primary studies were conducted in EMBASE, Medline and PubMed as listed in Table 13.

Table 13 Electronic databases searched for primary studies (scenario 3)

Database	Period covered
Embase.com (includes EMBASE and Medline)	2010–2012 (search performed on 3 December 2012)
PubMed	2010–2012 (search performed on 4 December 2012)

The search terms used to identify primary studies were adapted from the AHRQ 2011 search strategy and are listed in Appendix F.

Selection criteria

The evaluation team conducted three systematic reviews of the medical literature to address these research questions. After duplicate publications were excluded, citations were appraised to determine eligibility using the criteria listed in Table 14.

Studies of the highest level of evidence available (Table 16) for each question were included, that is, lower levels of evidence were excluded.

In the absence of RCTs (level II evidence) to directly assess the impact of these technologies on patient outcomes, this review included studies that compared the new technologies or strategies to conventional Pap test cytology with manual reading and reported on one or more of the following outcomes:

- diagnostic sensitivity and specificity, or the ratio of true positive (TP) to false positive (FP) findings, or the incremental rate of TPs, for detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS [adenocarcinoma in situ]) in women with a possible or definite HSIL cytology result
- changes in management, eg difference in colposcopy rates
- Patient outcomes:
    - quality of life—patient preference, satisfaction, psychological distress or anxiety
    - patient compliance
    - safety—adverse events, avoidance of unnecessary treatments
    - incidence of cervical cancer
    - overall survival
    - cervical cancer-specific mortality.

Additional outcomes extracted from studies that reported one or more of the above outcomes included unsatisfactory rates, positive predictive value (PPV) and accuracy at a test threshold of possible or definitive LSIL.

Outcomes were considered in terms of the hierarchy of Arbyn et al. (2009) (see "Benefits of screening", page 22).

An established and comprehensive model of screening, diagnosis and treatment of cervical cancer was also utilised to estimate the relative effectiveness and cost-effectiveness of incorporating these technologies into the Australian screening program, informed by evidence identified in this literature review.

Table 14 Inclusion/exclusion criteria for the identification of relevant studies

Characteristic	Criteria: Scenario 1	Criteria: Scenario 2	Criteria: Scenario 3
Study design	Inclusion: Systematic review, Clinical study or Australian evidence-based decision model Exclusion: Non-clinical study	Inclusion: Clinical studies Exclusion: Non-systematic reviews, letters, editorials, animal, in-vitro, laboratory studies, case studies, conference abstracts and technical reports, guidelines and systematic reviews that have been superseded	Inclusion: Clinical studies Exclusion: Non-systematic reviews, letters, editorials, animal, in-vitro, laboratory studies, case studies, conference abstracts and technical reports, guidelines and systematic reviews that have been superseded
Population	Primary screening population with data reported by age group	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Prior tests	Nil	Nil	Nil
Index test /Intervention	Cytology-based cervical screening program	Inclusion: Manual screening of LBC Automated cervical cytology image analysis for primary screening using LBC Exclusion: Technologies not commercially available Automated cervical cytology image analysis for primary screening of conventional Pap test cytology Automated cytology reading systems used for rescreening or quality control (QC)	Inclusion: HPV testing alone (1º question) HPV testing with cytology as either a co-test or a reflex test (2º questions) Exclusion: Technologies not commercially available[1]
Reference standard	N/A	Colposcopy with biopsy for positives	Colposcopy with biopsy for positives
Comparator	Inclusion: Start age: 25 years vs <25 years; Exit age: 65 years vs 70 years; Screening interval: 3-5 years vs 2 years or vs no screening	Conventional cytology (1º question)	Conventional cytology (1º question) HPV as a primary test (2º questions)
Outcomes[1]	Inclusion: CIN2, CIN3, cervical cancer, cervical cancer mortality, screening participation, adverse events Exclusion: Age-specific outcomes not reported	Inclusion: Accuracy (sensitivity and specificity, or the ratio of true positive to false positive findings, or the incremental rate of true positives) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS [adenocarcinoma in situ]) in women with a possible or definite HSIL cytology result. Relative accuracy measures if no reference standard applied to negatives. Absolute accuracy measures if reference standard applied to all or random selection of negatives Exclusion: Test yield or unsatisfactory rates without reported accuracy data	Inclusion: Cervical cancer incidence. Mortality. Accuracy (sensitivity and specificity, or the ratio of true positive to false positive findings, or the incremental rate of true positives) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS [adenocarcinoma in situ]) in women with a possible or definite HSIL cytology result. Relative accuracy measures if no reference standard applied to negatives Exclusion: Test yield or unsatisfactory rates without reported accuracy data
Publication type	English language	Inclusion: RCT or pseudorandomised trial (LBC manual) Exclusion: Non-English language. Superseded by AHRQ (Vesco et al. 2011) report Historical comparisons. Cohort studies (LBC manual)	Inclusion: RCT Exclusion: Non-English language. Superseded by AHRQ (Vesco et al. 2011) report

1. No studies were excluded based on HPV test type. The majority of studies used HC2. Some studies used GP5+/6+ PCR which is not commercially available but were not excluded as they provided a high level of evidence (RCT, level II) and were thus considered informative. Comparisons of the diagnostic performance of individual HPV tests is beyond the scope of this review but the potential for small differences to impact on overall program performance should be noted. For guidelines on assessing new HPV tests against HC2 see Meijer (2009).

Search results

Age range and screening interval

Existing HTA reports and systematic reviews

Since the IARC review, four additional HTAs and systematic reviews published between 2004 and November 2012 have assessed the impact of using different screening age ranges and intervals on screening program outcomes. Three of these reviews are included in the present assessment as the most recent high-quality English-language systematic reviews to assess these questions. These are:

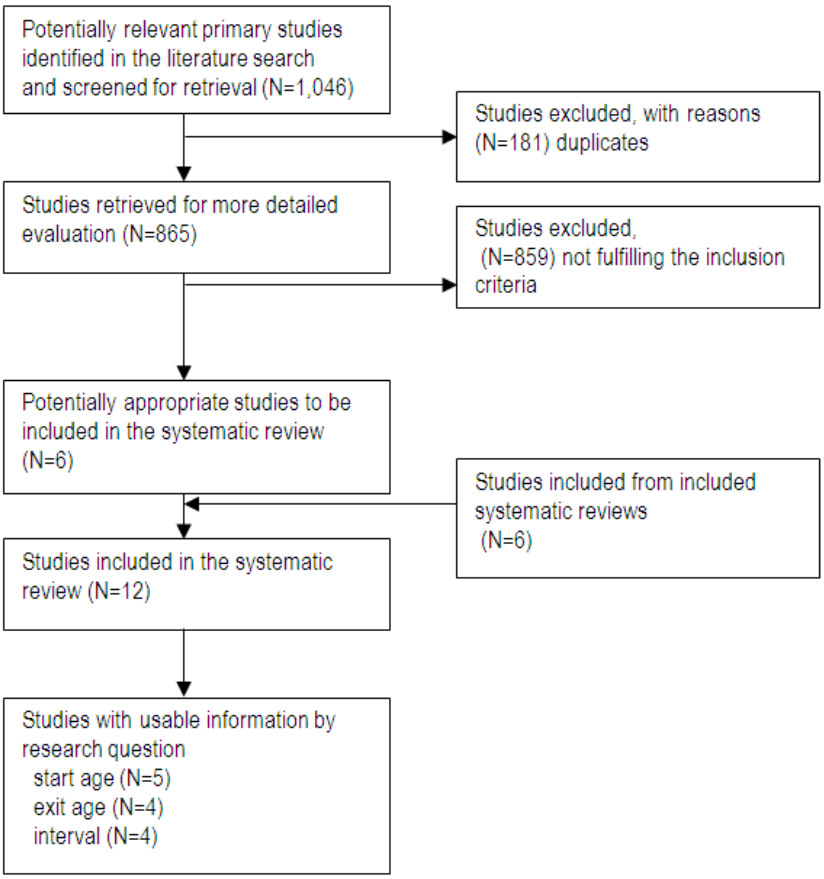
- the AHRQ review (Vesco et al. 2011) conducted to inform the US Preventive Services Task Force Recommendations for Cervical Screening
- The American Cancer Society (ACS), American Society for Colposcopy and Cervical Pathology (ASCCP), and American Society for Clinical Pathology (ASCP) review (Saslow et al. 2012) conducted to inform the ACS–ASCCP–ASCP guidelines for cervical screening the
- McMaster University Evidence Review and Synthesis Centre (ERSC) review (Peirson et al. 2012) conducted to inform the Canadian Task Force on Preventive Health Care Recommendations for Cervical Screening.

Primary studies

An additional six eligible studies were identified by updating the search conducted for the included AHRQ (Vesco et al. 2011) review (five primary studies and one Australian-modelled analysis).

Quorum flowchart

The process for identification of studies included in this review is depicted below (Figure 7).



Adapted from Moher et al. (1999)

Figure 7 Summary of the process used to identify and select primary studies examining screening age range and interval for research question 1 (scenario 1)

Liquid-based cytology (scenario 2)

Ten HTAs or systematic reviews relevant to the questions considering LBC or automated image analysis published between September 2008 and November 2012 were identified (see Appendix G, page 215).

The most recent high-quality systematic review identified that addressed the research question regarding the relative effectiveness of LBC and conventional cytology was Peirson et al. (2012). This systematic review searched for evidence on the incidence and mortality from cervical cancer following screening with liquid-based, versus conventional cytology techniques and also for automated-image analysis compared to conventional cytology. The AHRQ HTA (Vesco et al. 2011) also considered evidence on the relative accuracy of LBC and conventional cytology.

LBC

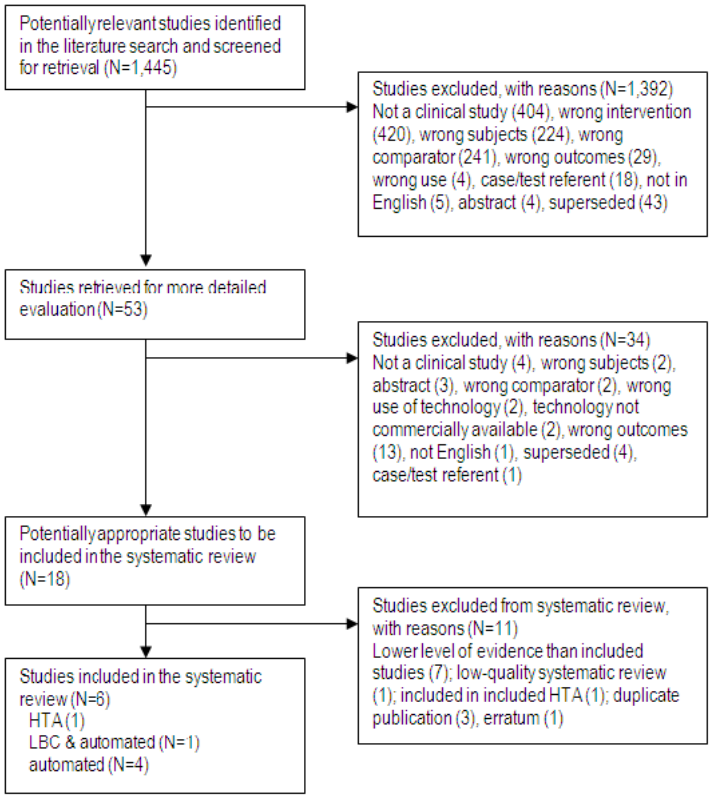
The Canadian systematic review (Peirson et al. 2012) did not include any studies that reported on the incidence and mortality from cervical cancer following screening with liquid-based, versus conventional cytology techniques. Thus the current review considers and updates evidence of these techniques from the AHRQ report which considers a broader range of outcomes. A systematic literature search for LBC primary studies was conducted from 2010 to update this report. In addition, the current reviewers re-assessed studies identified in the AHRQ search but excluded from that review.

Automated image analysis

The current review considers and updates evidence from the Canadian (Peirson et al. 2012) systematic review. As this review only included evidence for the outcomes of incidence and mortality, a systematic literature search for primary studies of automated image analysis technologies was conducted from 2008, to update the previous MSAC (2009) report.

Quorum flowchart

The results from the literature searches for LBC (2011–2012) and automated image analysis (2008–2012) were pooled. The process for identification of studies included in this review is depicted below (Figure 8).



Adapted from Moher et al. (1999)

Figure 8 Summary of the process used to identify and select primary studies examining LBC for research question 2 (scenario 2)

HPV testing (scenario 3)

Existing HTA reports and systematic reviews

The most recent high-quality systematic reviews identified that addressed the research question regarding the relative effectiveness of HPV and conventional cytology were:

- Canadian Task Force on Preventive Health Care systematic review (Peirson et al. 2012)
- Agency for Healthcare Research and Quality (AHRQ) HTA conducted for the US Preventive Services Task Force (USPSTF) (Vesco et al. 2011)
- Systematic reviews and meta-analyses by Arbyn (2012 and 2013).

The AHRQ review considered evidence to September 2010 and conducted an ancillary search which identified additional studies published to August 2011. This report considered evidence which addressed many of the research questions of the current review including:

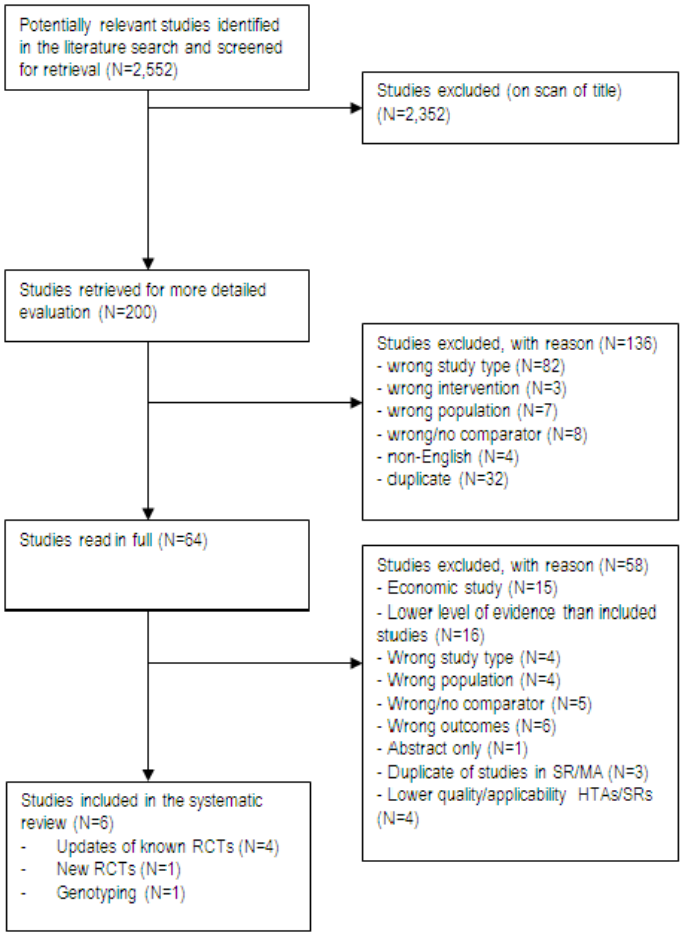
- HPV testing for primary screening, with and without cytology co-testing or triage of positive results
- Harms of HPV testing.

The Canadian Task Force on Preventive Health Care (CTFPHC) review (Peirson et al. 2012) considered evidence for primary or reflex HPV testing compared to conventional cytology to February 2011 (as a contextual (secondary) question). The review included studies reporting on the outcomes of incidence or mortality from cervical cancer. The review also included several questions on the harms of cervical cancer screening, including evidence of the value (acceptability, participation rates) of women's self-collection for HPV testing.

Arbyn (2012) considered accuracy outcomes.

Quorum flowchart

The process of the literature searches for HPV (2011–2012) for the identification of studies included in this review is depicted in Figure 9.



Adapted from Moher et al.(1999)

Figure 9 Summary of the process used to identify and select primary studies examining HPV for research question 3 (scenario 3)

Appraisal of the evidence

Appraisal of the evidence was conducted in three stages:

Stage 1: Appraisal of the applicability and quality of individual studies included in the review.

Stage 2: Appraisal of the precision, size and clinical importance of the primary outcomes used to determine the safety and effectiveness of the intervention.

Stage 3: Integration of this evidence for conclusions about the net clinical benefit of the intervention in the context of Australian clinical practice.

Validity assessment of individual studies

The evidence presented in the selected studies was assessed and classified using the dimensions of evidence defined by the National Health and Medical Research Council (NHMRC 2000).

These dimensions (Table 15) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from the literature identified as informing a particular intervention. The last two require expert clinical input as part of its determination.

Table 15 Evidence dimensions

Type of evidence	Definition
Strength of the evidence	
Level	The study design used, as an indicator of the degree to which bias has been eliminated by design.*
Quality	The methods used by investigators to minimise bias within a study design.
Statistical precision	The *p*-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect.
Size of effect	The distance of the study estimate from the "null" value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

* See Table 16

Strength of the evidence

The three sub-domains (level, quality and statistical precision) are collectively a measure of the strength of the evidence.

Level

The "level of evidence" reflects the effectiveness of a study design to answer a particular research question. Effectiveness is based on the probability that the design of the study has reduced or eliminated the impact of bias on the results.

The NHMRC evidence hierarchy provides a ranking of various study designs ('levels of evidence') by the type of research question being addressed (see Table 16).

Table 16 Designations of levels of evidence according to type of research question (NHMRC 2009)

Level	Intervention	Diagnostic accuracy	Prognosis	Aetiology	Screening intervention
I	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies	A systematic review of level II studies
II	A randomised controlled trial	A study of test accuracy with an independent, blinded comparison with a valid reference standard, among consecutive persons with a defined clinical presentation	A prospective cohort study	A prospective cohort study	A randomised controlled trial
III-1	A pseudo-randomised controlled trial (ie alternate allocation or some other method)	A study of test accuracy with an independent, blinded comparison with a valid reference standard, among non-consecutive persons with a defined clinical presentation	All or none	All or none[8]	A pseudo randomised controlled trial (ie alternate allocation or some other method)
III-2	A comparative study with concurrent controls: ▪ non-randomised, experimental trial ▪ cohort study ▪ case-control study ▪ interrupted time series with a control group	A comparison with reference standard that does not meet the criteria required for Level II and III-1 evidence	Analysis of prognostic factors among persons in a single arm of a randomised controlled trial	A retrospective cohort study	A comparative study with concurrent controls: ▪ non-randomised, experimental trial ▪ cohort study ▪ case control study
III-3	A comparative study without concurrent controls: ▪ historical control study ▪ two or more single-arm study ▪ interrupted time series without a parallel control group	Diagnostic case control study[6]	A retrospective cohort study	A case control study	A comparative study without concurrent controls: ▪ historical control study ▪ two or more single-arm study
IV	Case series with either post-test or pre-test/post-test outcomes	Study of diagnostic yield (no reference standard)	Case series, or cohort study of persons at different stages of disease	A cross-sectional study or case series	Case series

Table notes can be found in NHMRC 2009.

Individual studies assessing diagnostic effectiveness were graded according to pre-specified quality and applicability criteria (MSAC 2005), as shown in Table 17.

Table 17 Grading system used to rank included studies

Validity criteria	Description	Grading system
Appropriate comparison	Did the study evaluate a direct comparison of the index test strategy versus the comparator test strategy?	C1 direct comparison CX other comparison
Applicable population	Did the study evaluate the index test in a population that is representative of the subject characteristics (age and sex) and clinical setting (disease prevalence, disease severity, referral filter and sequence of tests) for the clinical indication of interest?	P1 applicable P2 limited P3 different population
Quality of study	Was the study designed to avoid bias? High quality = no potential for bias based on pre-defined key quality criteria Medium quality = some potential for bias in areas other than those pre-specified as key criteria Poor quality = poor reference standard and/or potential for bias based on key pre-specified criteria	Q1 high quality Q2 medium Q3 poor reference standard poor quality or insufficient information

Applicability

Studies of a population representing the Australian population with a mixture of patients presenting for screening or diagnosis were considered highly applicable. Studies of patients in a high-risk population with seeded cases or where different cytological preparation systems were pooled were considered of limited applicability. Studies in which all patients were undergoing cervical cytology for diagnostic purposes were considered not applicable. Although some studies have been conducted in a setting where the population is likely to be highly applicable, where this was not clearly reported the population will be rated as of limited applicability.

Quality

Seven criteria were applied to assess the quality of systematic reviews (Table 18). For the criterion addressing heterogeneity, systematic reviews that did not undertake a meta-analysis were rated 'not applicable' (N/A), unless heterogeneity was specifically mentioned. Studies were required to meet all seven criteria to be assessed as high quality. A study with four or fewer 'yes' or 'N/A' ratings was considered to be of low quality.

Table 18 Criteria used to assess the quality of effectiveness studies (adapted from NHMRC 2000 and Centre for Reviews and Dissemination 2001)

Study design	Quality checklist
Systematic reviews [a]	Was the research question specified?
	Was the search strategy explicit and comprehensive?
	Were the eligibility criteria explicit and appropriate?
	Was a quality assessment of included studies undertaken?
	Were the methods of the study appraisal reproducible?
	Were sources of heterogeneity explored?
	Was a summary of the main results clear and appropriate?

[a] High quality: Yes or N/A to all seven criteria; low quality: ≤4 Yes or N/A. Other studies assessed as fair quality.

Studies of diagnostic accuracy were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) quality assessment tool (Whiting 2003). Studies were required to meet all 12 criteria to be assessed as high quality (see details in footnote to Table 19), including a valid reference standard for at least all discordant positive slides (Davey et al. 2006). High-quality studies were required to use histology as a reference standard for comparisons of different slide preparation techniques. Adjudicated cytology reading was considered a valid reference standard for paired studies of alternate slide reading methods (Irwig et al. 2004).

Table 19 Criteria used to assess the quality of diagnostic accuracy studies—the QUADAS tool (adapted from Whiting et al. 2003)

Item	Criteria
1	Were patients prospectively recruited?
2	Were patients consecutively recruited?
3	Were selection criteria explicitly described?
4	Is the reference standard likely to correctly classify the target condition? Optimal (histology) / valid (consensus cytology) / invalid
5	Did all positive patients receive verification using a reference standard? All / positive / discordant participants
6	Is the time period between reference standard, comparator and index test short enough to be reasonably sure that the target condition did not change between the tests?
7	Was the test threshold specified?
8	Were test/comparator results interpreted blind to reference standard?
9	Were reference standard results interpreted blind to test/comparator results?
10	Were uninterpretable/intermediate test results reported?
11	Were withdrawals from the study explained?
12	Were sufficient data for determination of relative true and false positive rates reported?

High quality: Yes to 1, 3, 4, 5, 6, 10, 11; other items required to be either Yes or Unclear.
Low quality: No/unclear for 4 or 5; ≤ 4 yes or N/A ratings.
Other studies are assessed as fair quality.

Criteria for appraising the quality of RCTs of alternate screening strategies, with relative accuracy outcomes were adapted from the QUADAS tool (adapted from Whiting et al. 2003) with an additional two items considering random allocation, concealment and intention-to-screen analysis (from Centre for Reviews and Dissemination 2001). Thirteen criteria were applied to assess the quality of RCTs (see Table 20).

Table 20 Criteria used to assess the quality of screening RCTs with diagnostic accuracy outcomes—modified from the QUADAS tool (adapted from Whiting et al. 2003) and Centre for Reviews and Dissemination (2001)

Item	Criteria
1	Was assignment to the treatment groups really random?
2	Was allocation to intervention groups concealed from those responsible for recruiting the subjects?
3	Were patients consecutively recruited?
4	Were selection criteria explicitly described and appropriate?
5	Is the reference standard likely to correctly classify the target condition? Optimal (histology) / valid (consensus cytology) / invalid
6	Did a minimum of all patients testing positive receive verification using the same reference standard? All positive / X% positive / discordant participants
7	Is the time period between reference standard, comparator and index test short enough to be reasonably sure that the target condition did not change between the tests?
8	Was the test threshold specified?
9	Were test/comparator results interpreted blind to reference standard?
10	Were reference standard results interpreted blind to test/comparator results?
11	Were uninterpretable/intermediate test results reported?
12	Were withdrawals from the study explained and appropriate?
13	Were sufficient raw data for determination of relative true and false positive rates reported?

High quality: Yes to 1, 4, 5, 6, 8, 9, 11, 12, 13; other items required to be either Yes or Unclear.
Low quality: No/unclear for 5; ≤ 6 yes or N/A ratings.
Other studies are assessed as fair quality.

Statistical methods

Statistical precision was determined using statistical principles. Small confidence intervals and p-values give an indication as to the probability that the reported effect is real and not attributable to chance (NHMRC 2000). Studies need to be appropriately powered to ensure that a real difference between groups will be detected in the statistical analysis.

In a randomised setting, or where two tests are performed in the same subjects, the relative sensitivity (which can be estimated from the ratio of the detection rates) can be used as a valid diagnostic outcome (Arbyn et al. 2009). This outcome can be calculated from the available data in some studies that report absolute accuracy measures which are not valid due to incomplete verification of negative subjects. Thus the relative sensitivity between two tests was estimated from the ratio of detection rates of histologically confirmed high-grade lesions in RCTs (where the prevalence in each arm is considered equivalent).

$$Relative\ sensitivity\ = \frac{(TP\ index\ test\ /\ Number\ screened)}{(TP\ conventional\ cytology\ /\ Number\ screened)}$$

In studies with both tests performed in all subjects (eg split-sample), the number of true positive cases detected was compared between alternative testing strategies by $\chi 2$-test.

Estimated differences in sensitivity and specificity are calculated for studies with complete verification of discordant positive findings (at a minimum) comparing two tests in the same population. While the total number of cases is not known, and absolute sensitivity and specificity are not known, the difference can be estimated from the known values as below:

$$Sensitivity\ difference\ = \frac{Additional\ TP}{(TP\ +\ Known\ FN\ cases)}$$

$$Specificity\ difference\ = \frac{Fewer\ FP\ cases}{(Concordant\ negative\ +\ FP\ cases)}$$

Data from RCTs of HPV screening strategies were analysed following the methods described in the AHRQ HTA (Vesco et al. 2011, p. 122). The relative detection rate (also termed the 'relative sensitivity') was calculated and where possible, the relative false positive proportion and relative positive predictive value. Results were analysed by women screened rather than women randomised. For the calculation of relative false positive proportion and relative positive predictive value, the lowest referral criterion to define a positive screening test was used. This may overestimate false positives for trials with higher initial referral criteria but relative test performance measures may be less affected.

Size of effect

For intervention studies it was important to assess whether statistically significant differences between the comparators were also clinically important. The size of the effect needed to be determined, as well as whether the 95% confidence interval included only clinically important effects.

Relevance of evidence

The outcomes being measured in this report should be appropriate and clinically relevant. Inadequately validated (predictive) surrogate measures of a clinically relevant outcome should be avoided (NHMRC 2000).

The use of precancerous cervical lesions as a surrogate for cervical cancer is justified by existing evidence that early detection and treatment of precancerous cervical lesions leads to a reduction in the incidence and mortality of cervical cancer (AIHW 2012; Peto et al. 2004b, See 'Cervical cancer trends in Australia', page 10). Carcinoma in situ is the most certain surrogate for invasive cervical cancer (Schiffman et al. 2007). CIN2 lesions can be caused by HPV types rarely found in cancer and have a sizeable regression potential whereas CIN3 lesions share the same HPV type spectrum and causal cofactors as cancer, although those detected by screening may be small and not certain to progress. The use of precancerous surrogates in test performance studies and trials is a limitation of the evidence and the degree of benefit in preventing invasive cervical cancer is usually uncertain.

Assessment of the body of evidence

Appraisal of the body of evidence was conducted along the lines suggested by the NHMRC in their guidance on clinical practice guideline development (NHMRC 2008). Five components are considered essential by the NHMRC when judging the body of evidence:

- the evidence base—which includes the number of studies sorted by their methodological quality and relevance to patients
- the consistency of the study results—whether the better quality studies had results of a similar magnitude and in the same direction, ie homogenous or heterogenous findings

- the potential clinical impact—appraisal of the precision, size and clinical importance or relevance of the primary outcomes used to determine the safety and effectiveness of the test

- the generalisability of the evidence to the target population

- the applicability of the evidence—integration of this evidence for conclusions about the net clinical benefit of the intervention in the context of Australian clinical practice.

A matrix for assessing the body of evidence for research questions, according to the components above, was used for this assessment (see Table 21) (NHMRC 2008).

Table 21 Body of evidence assessment matrix

	Body of evidence			
Component	A Excellent	B Good	C Satisfactory	D Poor
Evidence base	several level I or II studies with low risk of bias	one or two level II studies with low risk of bias or a SR/multiple level III studies with low risk of bias	level III studies with low risk of bias, or level I or II studies with moderate risk of bias	level IV studies, or level I to III studies with high risk of bias
Consistency	all studies consistent	most studies consistent and inconsistency may be explained	some inconsistency reflecting genuine uncertainty around clinical question	evidence is inconsistent
Clinical impact	very large	substantial	moderate	slight or restricted
Generalisability	population/s studied in body of evidence are the same as the target population	population/s studied in the body of evidence are similar to the target population	population/s studied in body of evidence different to target population for guideline but it is clinically sensible to apply this evidence to target population	population/s studied in body of evidence different to target population and hard to judge whether it is sensible to generalise to target population
Applicability	directly applicable to Australian healthcare context	applicable to Australian healthcare context with few caveats	probably applicable to Australian healthcare context with some caveats	not applicable to Australian healthcare context

Adapted from NHMRC (2008)

Expert advice

The National Cervical Cancer Renewal Steering Committee provided guidance to the health technology assessors to ensure that the assessment is clinically relevant and takes into account consumer interests. Membership of the Committee is provided at Appendix B.

Results of assessment—Safety

Conventional cytology (scenario 1)

Summary 1: Summary of safety—conventional cytology (scenario 1)

Narrowing the screening age range and extending the screening interval will provide potential safety benefits by reducing the harms of tests and procedures for abnormalities that may otherwise regress. Estimates of these benefits need to be weighed against uncertainty about the potential impact on cervical cancer incidence and mortality.

The key safety concern about narrowing the age range and broadening the interval for screening to meet IARC recommendations is the potential for an increase in cervical cancer incidence and mortality due to the missed opportunity for earlier detection and treatment of precancerous lesions and invasive cancers. Narrowing the age range, two age groups of women would no longer be invited for screening: (i) women younger than 25 years of age; and (ii) women aged 66–70 years who have participated in the NCSP with at least two consecutive prior negative screening tests. Broadening the screening interval will potentially reduce the opportunity for earlier detection of previously undetected and new abnormalities in asymptomatic women who have tested negative on the prior screening test—with potential delays in detection of up to one year for women aged 25–50 years if the screening interval is extended to three years; and up to three years for women aged 25–65 years if the screening interval is extended to five years. The potential for harm from these changes will depend on the risk of precancerous lesions and invasive cancers in each age group and the consequences of delayed diagnosis.

There are also potential safety benefits for adopting a less intensive screening strategy as outlined on page 23. These benefits include reduced patient discomfort due to reduced testing, reduced anxiety due to the detection of cytological abnormalities that are subsequently found to be clinically inconsequential (biopsy finding <CIN2); and reduced risk of treatment and reproductive complications for CIN2/3 lesions that may subsequently regress if left undiagnosed. Potential harms would therefore need to be weighed against these potential benefits.

Evidence from clinical studies about the impact of changing the screening age range and interval on cervical cancer incidence and mortality is summarised below in the effectiveness section (page 59).

A comprehensive mathematical model of screening, diagnosis and treatment was developed to provide estimates of patient outcomes and resource use of screening using the IARC age range and screening interval versus the current NCSP strategy for the assessment of the safety, effectiveness and cost-effectiveness of this strategy in the Australian context.

Liquid-based cytology (scenario 2)

Summary 2: Summary of safety—liquid-based cytology (scenario 2)

LBC with manual or automated slide reading is considered a safe procedure.

LBC with manual or automated slide reading uses the same procedure for collecting cervical cell samples as conventional Pap cytology tests and therefore does not introduce any additional risks to the patient. Collection of cervical cells is regarded as safe. Some women may experience discomfort or minor bleeding afterwards that resolves spontaneously.

Safety of screening with LBC must consider any change in the false positive rate and the downstream effects of this.

HPV testing (scenario 3)

Summary 3: Summary of safety— HPV testing (scenario 3)

HPV testing is a safe procedure.

HPV testing uses the same procedure for collecting cervical cell samples as used in conventional Pap cytology tests and therefore does not introduce any additional risks to the patient. Collection of cervical cells is regarded as a safe procedure. Some women may experience discomfort or minor bleeding afterwards that resolves spontaneously.

No adverse events due to the collection of cervical cells for HPV testing or cytology were reported in the seven trials included in this review (327,187 patients).

Safety of screening with HPV testing must consider any change in the false positive rate and the downstream effects of this (additional investigation and monitoring and associated patient anxiety).

An additional consideration is whether HPV testing is associated with any psychological harm. The AHRQ (Vesco et al. 2011) reviewed this and their conclusions are summarised below. Note that one of the included studies is an Australian RCT comparing HPV testing with a repeat smear for triaging women with minor cytological abnormalities (McCaffery 2010).

Harms of HPV testing (AHRQ executive summary, Vesco et al. 2011)

In addition to concerns about false positive testing and related harms, we identified four studies that described potential psychological harm from HPV testing. In the short term (first few weeks after receiving test results), women who test positive for HPV had higher levels of anxiety and distress and greater concerns about their health and health risks. In the long term, however, these results did not persist. In fact, when considering triage of ASC-US cytology with an HPV test versus repeat cytology, long-term follow-up suggests greater satisfaction with care and less distress among women undergoing HPV testing. This may be because women who undergo repeat cytology have to wait for additional results before it is determined whether or not they need colposcopy, whereas women undergoing HPV testing are triaged much more quickly.

The evidence about harms of HPV testing is limited. Only two of the four included studies present long-term follow-up; there was a small number of women included in the follow-up; only one study administered questionnaires prior to cytology and HPV testing; and all studies had large proportions of women who did not return the study questionnaires. Larger studies with longer term follow-up, assessment of psychological measures pre- and post-test, and adjustment for baseline psychological measures and appropriate confounders are needed to determine the psychological impact of HPV testing.

Results of assessment—Effectiveness: Conventional cytology (scenario 1)

Primary research question 1

What is the comparative safety and effectiveness of conventional cytology, using the IARC recommendations for age range and interval, compared with the protocol used in the current Australian cervical screening program?

Summary 4: Summary of effectiveness—cytology-based cervical screening

Start age 25 years versus 20 years

Cervical cancer is very rare before the age of 25 years and cervical cancer mortality rates in women aged 20–24 years have not changed since the introduction of the national screening program. However, screening women below the age of 25 is more likely to lead to further investigation than screening for older women.

In Australia, NCSP data show that among women aged 20–69 years, women aged 20–24 years have the highest risk of abnormal cytology results, the second highest risk of high-grade histology after women aged 25–29 years, but the lowest risk of cervical cancer. These data indicate that screening women aged 20–24 years is more likely to lead to further investigation and potentially treatment than older age groups of women, despite their low risk of cancer.

Changing the age for cervical screening to commence at 25 years will reduce the rate of investigation in women aged less than 25 years. Observational studies indicate that most HPV infection and associated minor cervical changes will regress; however, some women will have persistent infection with CIN present when screening commences at age 25 years. This will increase the incidence of pre-invasive changes detected in unvaccinated women aged 25 years and older.

One large UK case control study has examined the impact of screening in women aged 20–24 years on cervical cancer incidence (Sasieni et al. 2009). This study found that screening women under the age of 25 years does not reduce the incidence of cervical cancer at ages 25–29 years.

Following an increase in the commencement age for screening in England from 20 to 25 years in 2004, an increase in the incidence of cervical cancer and high-grade CIN in women aged 25–29 years was observed. However, these cancer incidence trends were also observed in Scotland and Wales which continued to start screening at age 20 years, indicating the impact of factors other than screening age (Sasieni & Castanon 2012).

The national HPV vaccination program is anticipated to significantly reduce the risk of cervical cancer in young women. The early Australian data show a reduction in high-grade abnormalities.

Exit age 65 years versus 69 years

No studies were identified that directly compared the effectiveness of exiting screening at 65 versus 69 years of age.

Two case control studies provide evidence that screening beyond 65 years can reduce the risk of cervical cancer (Andrae et al. 2008; Lonnberg et al. 2012). However, neither of these studies reported on prior screening history to examine whether screening after 65 years offers any additional protection to that provided by adequate screening up to the age of 65 years.

A US health insurance plan member retrospective case series reported that 75% of women aged 65 and over that developed invasive cervical cancer did not meet their exit test criteria of three consecutive negative cytology test or a single HPV co-test (Dinkelspiel et al. 2012).

Screening interval three years versus two years in women aged under 50 years; and five years versus two years in women aged 50 years and older

The current Australian screening program is considered intensive in comparison to those of other developed countries with successful cervical cancer prevention programs.

A UK case control study provides evidence that for women under 40 years of age, the risk reduction offered by screening is diminished substantially for screening intervals greater than three years (Sasieni et al. 2003). In contrast, for older women, screening intervals of five years offer risk reductions approaching those observed at three years (Sasieni et al. 2003).

A modelling study of the Australian NCSP predicts that extending the screening interval to three years will not substantially alter cervical cancer incidence or mortality rates but will lead to a reduction in the number of cytology tests and colposcopy procedures (Creighton et al. 2010). These results are consistent with results from a US-modelled analysis (Kulasingam et al. 2011). They are based on data from unvaccinated women.

Conclusions

- Cervical cancer is very rare before the age of 25 years.
- HPV vaccination is anticipated to substantially reduce the risk of cervical cancer in young women.
- There are limited comparative data on the age at which to start and stop screening.
- The available data do not demonstrate a benefit of screening women aged less than 25 years. However, screening in this age group does result in increased investigations, treatments and potential harms, without decreasing mortality.
- The majority of cervical cancer cases in women aged 65 and over are in women not meeting the criteria for an exit test.
- Extending the screening interval from two to three years is unlikely to substantially alter cervical cancer incidence or mortality rates but will lead to a reduction in the number of cytology tests and colposcopy procedures (based on two modelling studies).
- For women aged over 50 years screening intervals of five years offer risk reductions approaching those observed at three years in younger women (based on an analysis of cancer in women aged 55 to 69 years).

Table 22 Inclusion criteria for identification of studies examining screening age range and interval for research question 1 (scenario 1)

Research question:		
Selection criteria	Inclusion	Exclusion
Study design	Systematic review, clinical study or Australian evidence-based decision model	Non-clinical study
Population	Primary screening population with data reported by age group	-
Prior tests	Nil	-
Index test/Intervention	Cytology-based cervical screening program	-
Reference standard	N/A	-
Comparator	Start age: 25 years vs <25 years; and/or Exit age: 65 years vs 70 years; and/or Screening interval: 3–5 year vs 2 year or vs no screening	-
Outcomes	CIN2, CIN3, cervical cancer, cervical cancer mortality, screening participation, adverse events	Age-specific outcomes not reported
Publication type	English language	-

Included studies

The inclusion criteria used for identification and inclusion for the assessment of evidence for screening age range and interval are listed in Table 22.

No studies were identified that compared a cervical screening program strategy using the IARC recommendations for age range and interval, versus the current NCSP strategy used in Australia, to provide direct evidence about the impact of adopting the IARC recommendations on cervical cancer incidence or mortality.

Therefore, this review presents the evidence available to address individual sub-questions about the impact of changing the screening start age (25 years versus 20 years), exit age (65 versus 70 years) or screening interval (three or five years versus two years) on cervical incidence and mortality.

The literature review identified three recent systematic reviews (Vesco et al. 2011; Saslow et al. 2012; Peirson et al. 2012). These reviews each investigated different questions about the optimal age range and/or screening interval of cytology-based cervical screening programs for inclusion in the present review.

From these reviews, six primary studies were identified that addressed one or more sub-questions. An additional five primary studies were identified from the literature search conducted to update the published systematic reviews (Sasieni & Castanon 2012 Patel et al. 2012; Dinkelspiel et al. 2012; Lonnberg et al. 2012 & 2013). An Australian modelling study that examined the impact of changing the screening interval was also included (Creighton et al. 2010).

In addition to this research evidence, the following Australian data are also considered:

- AIHW/Australian Cancer Registries data on temporal trends in age-specific cervical cancer incidence and mortality rates (Figure 4 and Figure 5, page 11)

- NCSP data on age-specific risks of abnormal cytology findings and high-grade histology are reported to examine differences in the ratio of abnormal cytology findings to cancer detection rates across different age groups (Table 8 and Table 9, page 33).

Study characteristics, quality and applicability

Systematic reviews and HTA

The Peirson et al. (2012) review conducted for the CPHTFC provides the most comprehensive and up-to-date assessment of the evidence relevant to scenario 1 for the present review. It includes an assessment of screening start and exit ages, and screening interval, with a literature search conducted up to April 2012. It also includes an assessment of screening harms in women under the age of 30 years to inform decisions about screening start age (see summary of treatment harms page 23). The AHRQ (Vesco et al. 2011) review conducted for the USPSTF assessed the start age for screening with a literature search conducted up to August 2011; and the Saslow et al. (2012) review conducted for the American Cancer Society (ACS), the American Society for Colposcopy and Cervical Pathology (ASCCP), and the American Society for Clinical Pathology (ASCP) assessed the exit age, and the optimal screening intervals for women aged 21–29 years and 30 years and older, with a literature search conducted up to July 2011.

In addition to these systematic reviews, the USPSTF contracted a decision model analysis to consider the impact of screening if the age range was altered between a start age from 15 to 25 years; and an exit age from between 65 to 90 years; and the screening interval was altered between one to five years, using colposcopies per life-year as the primary outcome to capture the trade-off between the benefits of screening versus the potential for harm from additional investigations initiated by screening (Kulasingam 2011).

All three reviews met the criteria for a high-quality systematic review. The review questions, methods and the corresponding guideline recommendations are summarised in Table 23.

Table 23 Systematic reviews examining age range and interval for cytology screening: review questions, methods, quality and corresponding guideline recommendation

Author, year	Review questions : Age range Screening interval	Search methods	Evidence & Recommendation	Quality
AHRQ Vesco et al. 2011 Informed USPSTF United States	*Start age* • When should screening begin? • Should start age vary by test type, age, sexual history or other patient characteristics? *Exit age* Addressed in previou AHRQ systematic review (Hartmann et al. 2002) *Screening interval* Not addressed in systematic review	*Search period*: 2000–August 2011. *Sources:* HTA, Systematic reviews databases (2000–January 2007) Cochrane Central Trials Registry MEDLINE *Included studies* Systematic reviews RCTs Cohort studies Case control studies Ecological reports correlating population-based CIN rates and cancer detection with screening Decision model	*Start age* ▪Relative effectiveness of screening program by age group (2 studies) ▪Natural history of CIN and HPV infection in young women (2 studies) ▪Difference in risk of false positive findings by age group (1 study) *USPSTF 2012 Recommendation* Start age: 21 years Exit age: 65 years if adequate prior screening Screening interval cytology every 3 years, or for women ≥30 years; HPV-cytology co-testing every 5 years For exit age 65 years, adequate prior screening = 3 consecutive negative cytology tests or 2 consecutive negative HPV tests within 10 years, including one within 5 years. *Note*: The USPSTF also commissioned a decision model to address questions about start age, exit age and screening interval (Kulasingam 2011).	*Quality:* High Explicit review questions: yes Explicit and appropriate eligibility criteria: yes Explicit and comprehensive search strategy: yes Quality of included studies appraised: yes Methods of study appraisal reproducible: yes Heterogeneity between studies assessed: yes Summary of main results clear and appropriate: yes
ERSC Peirson et al. 2012 Informed CTFPHC Canada	*Start age* • What is the effect of varying start age on incidence and mortality of invasive cervical cancer? • What are the harms of screening for women <30yrs? *Exit age* • What is the effect of varying exit age on incidence and mortality of invasive cervical cancer *Screening interval* • What is the effect of varying the screening interval on incidence of invasive cervical cancer?	*Search period:* 1995–April 2012 *Sources:* Cochrane Central Trials Registry MEDLINE, EMBASE *Included studies:* Systematic reviews RCTs Cohort studies Case control studies Cancer registry incidence and mortality data pre- and post-introduction of screening	*Start age* ▪Relative effectiveness of screening program by age group (4 studies, including 1 study reporting separate data for women < 25 year) ▪ In women < 30 yr: Rate of abnormal Pap cytology, colposcopy and treatment (Canadian screening program data) -Treatment and associated harms (4 studies) *Exit age* ▪Relative effectiveness of screening program by age group (1 study reporting data for women >65 yr) *Screening interval* ▪Relative effectiveness of screening by interval length (14 studies, including 2 studies reporting data for age groups specified for present review) *CTFPHC 2013 Recommendation:* Start age: 25 years	*Quality:* High Explicit review questions: yes Explicit and appropriate eligibility criteria: yes Explicit and comprehensive search strategy: yes Quality of included studies appraised: yes Methods of study appraisal reproducible: yes Heterogeneity between studies assessed: yes Summary of main results clear and appropriate: yes

Author, year	Review questions : Age range Screening interval	Search methods	Evidence & Recommendation	Quality
			Exit age: 70 years if three negative tests in the last 10 years Screening interval: every 3 years	
Saslow et al. 2012 Informed ACS ASCCP ASCP United States	*Start age* Not addressed *Exit age* • Should screening be recommended for women ≥65 years with adequate negative prior screening? *Screening interval* • What are optimal screening intervals for cytology-based screening (1,2,3, 5 yrs) -for age 21–29 yrs -for age ≥30 yrs • Can screening interval be extended for these age groups after 2 negative screening tests?	*Search period:* 1995–July 2011 *Sources:* PubMed *Included studies:* RCTs Observational studies ≥100 subjects Decision models that reported effectiveness outcomes	*Exit age* ▪Relative effectiveness of screening women ≥65 yrs (1 modelling study) *Screening interval* ▪Relative effectiveness of screening by interval length (18 studies) .- impact of prior negative screening history on screening interval (2 studies) *ACS Recommendation:* Start age: 21 yr Exit age: 65 yr if adequate prior screening Screening interval: women 21–29 yrs cytology every 3 years women 30–65 yrs HPV-cytology co-testing (preferred) or cytology every 5 years (acceptable) For exit age 65 yr, adequate prior screening = 3 consecutive negative cytology tests or 2 consecutive negative HPV tests within 10 years, including one within 5 years.	*Quality:* Unclear Explicit review questions: yes Explicit and appropriate eligibility criteria: yes Explicit and comprehensive search strategy: yes Quality of included studies appraised: not clear Methods of study appraisal reproducible: no Heterogeneity between studies assessed: yes Summary of main results clear and appropriate: yes

ACS = American Cancer Society; ASCCP = American Society for Colposcopy and Cervical Pathology; ASCP= American Society for Clinical Pathology; CTFPHC= Canadian Task Force on Preventive Health Care; ERSC=Evidence Review Synthesis Centre; RCT= randomised controlled trial; USPSTF= US Preventive Services Task Force

Primary studies

Five studies reported data to examine the impact of changing the cytology screening start age on risk of cervical cancer (four studies); or risk of further investigation for women screened at 20–24 years versus 25–29 years (one study). No studies were designed to directly compare the impact of exiting screening at age 65 versus 70 years on cervical incidence or mortality. However, four studies reported data on cervical cancer incidence (three studies) or mortality (one study) and screening history in women over age 65 years to inform conclusions about potential impact of lowering the exit age. Three eligible studies reported data to compare the effectiveness of screening at different screening intervals. None of the included studies were conducted in a population vaccinated against HPV infection. Therefore, the prevalence and incidence of cervical cancer and precancerous cervical lesions in these studies will differ to that expected in Australia in the future, particularly in young women. The characteristics and quality appraisal of these studies are summarised in Table 24, Table 25 and Table 26.

Screening start age

One large UK case control study (Sasieni et al. 2009) and three population-based temporal trends studies examined the impact of changing the cytology screening start age from 20 to 25 years. Sasieni & Castanon (2012) and Patel et al. (2012) examined cervical cancer incidence in the UK before and after the screening start age was raised from 20 to 25 years in 2004. Sigurdsson (2010) examined cervical cancer incidence in Iceland before and after 1998 when the screening start age was lowered from 25 to 20 years (Sigurdsson 2010).

The Sasieni et al. (2009) case control study and the Sasieni & Castanon (2012) and Patel et al. (2012) population-based temporal trends studies are classified as providing NHMRC level III-2 evidence, with the latter two studies including a concurrent control group. The Sigurdsson et al. (2010) temporal trends study only included a historical control group and is therefore classified as providing NHMRC level III-3 evidence. The Sasieni et al. (2009) study is classified as providing higher quality evidence to examine the effectiveness of screening in women aged 20–25 years than the temporal trends studies because it compared the risk of cervical cancer by screening history in women drawn from the same source population, with prospectively recorded screening data available for cases and all randomly selected controls, which avoids recall bias and minimises selection bias. In contrast, the two UK temporal trends studies compared the risk of cervical cancer after the screening age was raised from 20 to 25 years in England with the corresponding changes over time in a control group from Wales (Patel et al. 2012), or Wales and Scotland (Sasieni & Castanon 2012), where the screening start age remained at 20 years. Concurrent comparisons across different regions have a greater potential for selection bias because different populations may have a different risk profile. The Sigurdsson (2010) study is classified as lower quality because of the high risk of confounding when relying on historical comparisons.

The Patel et al. (2012) study examined cervical cancer incidence between 2000 and 2009 to assess rates before and immediately after the screening age was raised to 25 years in England in 2004. The investigators included cases from a postcode defined region in the north-east of England (152 cancers in women aged <30 years). The Sasieni & Castanon (2012) study included cases reported England-wide from 2000–10, and further examined screening audit data available for young women between 2007 and 2011 (1,418 cancers in women aged <30 years between 2007 and 2011). This study was designed to investigate whether the increase in cervical cancer incidence observed in England in 2009 could be attributed to raising the screening start age from 20 to 25 years in 2004. The authors noted that women under 25 years who had entered the screening program before 2004 would have continued to be invited for three-yearly screening, and it would only be in 2009 that women aged 20 years in 2004 would enter the program for the first time at age 25 years.

Although assessed as higher quality compared to temporal trends studies, the Sasieni at al. study (2009) also has limitations, in particular the potential for selection bias due to possible differences in risk factors for cervical cancer between women participating in organised screening versus non-attendees. However, as the investigators have discussed, any such confounding factors associated with self-selection would need to work in different directions for different screening age groups to explain the different screening effects observed in different age groups (Sasieni et al. 2009).

These four studies were conducted in unvaccinated women. Applicability of the study results to vaccinated young women is not known. The risk of cervical cancer and therefore the impact of screening is expected to be lower in vaccinated women who will have a lower risk of HPV and precancerous changes.

One additional study reported on the risk of false positive results for women screened at 20–24 years versus 25–29 years (Insinga et al. 2004). This observational cohort study is classified as providing NHMRC level III-2 evidence about the difference in false positive screening results by age group with a low risk of bias. The study population included all women with abnormal cytology results from a sample of women enrolled in a US health insurance plan over two years. Due to potential differences in HPV and cervical cancer risk of US health plan enrollees and the general Australian screening population, the study results are judged as having fair applicability to the Australian setting. However, applicability of results to young vaccinated women is uncertain.

Table 24 Characteristics and quality of primary studies examining screening start age for cytology screening

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality and applicability
Sasieni et al. 2009 UK Participating health authorities 1990–2008 Updates Sasieni et al. 1996, 2003	N=11,901 • 4,012 invasive cancer Includes 437 cancers diagnosed <30 yr • 7,889 control • Screened and unscreened Age • 20–69 yr • Subgroups 20–24 yr 25–29 yr 30–34 yr ..... to 65–69 yr	Start age Relative effectiveness of screening vs no screening on cervical cancer incidence for comparison of 3-yr screening age groups, from: 20–22 yr to 62–64 yr (data not presented for comparison of screening vs no screening >65 yr)	Study design: Case control study Analysis Risk of cervical cancer • by screening status: Screened in last 3 yrs vs no screening in last 5 yrs • stratified by 5-yr age groups Odds ratio (95% confidence interval)	AHRQ Observational study NHMRC level III-2 Quality: Low risk of bias Applicable: Uncertain in vaccinated young women • Organised cytology screening program • Screening 20–24 vs 25–29 yr • Assessed 3-yr screening interval • Unvaccinated young women
Sasieni & Castanon 2012 UK • England • Wales • Scotland Cancer registry data 2000–2010 National audit data England 2007–2011	N= total NR Includes From England 2007–2011: • 4,079 cancers in women 20–39 yr • 1,418 cancers in women <30 yr Age • 20–50 yr • Subgroups 20–24 yr 25–29 yr 30–34 yr 35–39 yr	Start age Relative effectiveness of screening program with start age 20 yr vs 25 yr on cervical cancer incidence • Regional comparison England vs Wales vs Scotland England start age 25 yr Wales start age 20 yrs	Study design: Population-based Temporal trends with contemporaneous control group Analysis All cervical cancers • annual incidence 2000–2010, England • stratified by 5-yr age groups • with comparison to Wales, Scotland Risk ratio (95% confidence interval), test of trend Test of difference in trend for England vs Wales and Scotland between 2000–-2010 stratified by age group Test of trend for annual incidence cancers 2007-–2011, England • for age groups: 20–22, 23, 24, 25, 26, 27–29, 30–34, 35+ yr CIN3 • Annual incidence rates, 2003–2010, England	Search update Observational study NHMRC level III-2 Quality: Moderate risk of bias Applicable: Uncertain in vaccinated young women • Organised cytology screening program • +/– screening 20–24 yrs • 3-yr screening interval • Unvaccinated young women

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality and applicability
			• stratified by 5-year age groups	
Patel et al. 2012 UK • England north-east (NE) postcode defined region • Wales cancer and screening registries 2000–2009	N= total NR • Cancer 1,061 Includes 152 cancers diagnosed <30 yr • HG-CIN NR Age • 20–65+ yr • Subgroups 20–24 yr 25–29 yr 30–34 yr ≥ 35 yr	Start age Relative effectiveness of screening program with start age 20 yr vs 25 yr on cervical cancer incidence • Historical comparison Pre- vs post-2004 NE England start age 20 yr up to 2004 25 yr after 2004 • Regional comparison NE England vs Wales NE England start age 25 yr Wales start age 20 yr	Study design: Population-based Temporal trends with contemporaneous control group Analysis Incidence of cervical cancer NE England All cancers, Stage IB+ cancers • pre- vs post-2004 • stratified by 5-yr age groups • with comparison to Wales Incidence rate ratio (95% confidence interval) Interaction between relative risk of cervical cancer for young (20–24 yr, 25–29 yr) vs 30+ yr age groups and time period pre vs post 2004 Average annual % change (AAPC) in 3-year average incidence rate by age group • NE England and Wales Incidence of HG-CIN (as above) STD incidence • chlamydia, herpes simplex, genital warts, gonorrhoea, syphilis	Search update Observational study NHMRC level III-2 Quality: Moderate risk of bias Applicable: Uncertain in vaccinated young women • Organised cytology screening program • +/– screening 20–24 yr • 3-year screening interval • Unvaccinated young women
Sigurdsson 2010 Iceland Cancer screening registry 1964–2008 Updates Sigurdsson & Sigvaldason 2007	N=320,000 source population • CIN2 • CIN3 • Cancer Includes 70 cancers diagnosed < 30 yr Age • 20-69 yr • Subgroups	Start age Relative effectiveness of screening program with start age 20 yr vs 25 yr on cervical cancer incidence • Historical comparison Pre vs post 1998 Screening start age 25 yr up to 1998 20 yr after 1998	Study design: Population-based registry cohort Temporal trends with historical control group Analysis Incidence of cervical cancer All cancers, Stage IA, IB, IIA+ • Pre- vs post-1998 • stratified by 5-yr age groups Statistical test for difference in age-specific incidence rate by time period	AHRQ Observational study NHMRC level III-3 Quality: High risk of bias No contemporaneous comparison Applicable: Uncertain in vaccinated young women • Organised cytology screening program

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality and applicability
	20–24 yr 25–29 yr 30–34 yr 35–39 yr			• +/– screening 20–24 yrs • 2-year screening interval • Unvaccinated young women
Insinga et al. 2004 United States Health insurance members 1997–2002	N= 30,936 Women with cytology tests conducted in 1998 Age • 15+ yr • Subgroups 15–19 yr 20–24 yr 25–29 yr Then 10-yr age groups to 70–79 yr 80+ yr	Start age Incidence of false positive cytology results by age group	Study design: Health insurance member cohort study Analysis Age-specific rates of cytology abnormalities, CIN 1-3+, incomplete follow-up and false positive cytology results (no CIN/cancer) per 1000 women/ including 20–24 yr, 25–29 yr age groups	AHRQ Observational study NHMRC level III-2 Quality: Low risk of bias Applicable: Uncertain in vaccinated young women • US health plan enrollees

Screening exit age

The four studies that have examined the impact of screening in women over the age of 65 years include a case control study from Sweden that estimated the effectiveness of screening in reducing the risk of a cervical cancer diagnosis after the age of 65 years (Andrae et al. 2008), and two case control studies from Finland that estimated the effectiveness of screening in reducing the risk of a cervical cancer diagnosis in five-year age groups between 25 and 69 years (Lonnberg et al. 2012) and reducing the risk of cervical cancer mortality after the age of 65 years (Lonnberg et al. 2013) with correction for self-selection bias of screening participation. The fourth study was a retrospective analysis of cervical cancer incidence rates in women aged 65 years and older from a US health plan member cohort that reported the proportion of cancers that were detected in patients who had met criteria for exiting screening (Dinkelspiel et al. 2012).

The Andrae et al. (2008) study is classified as providing NHMRC level III-2 evidence of the effectiveness of screening. This study defined exposure to screening as those complying with screening between 6 and 66 months prior to diagnosis. The study has a moderate risk of bias due to excluding screening undertaken within the six months prior to the cancer diagnosis from the definition of screening exposure, which may overestimate the effect of screening.

The Lonnberg et al. (2012 & 2013) studies are classified as providing NHMRC level III-2 evidence with low risk of bias due to inclusion of cases from the national cancer registry and population register sampling of controls to avoid selection bias; and the use of prospectively recorded screening program data to accurately identify past screening history to avoid recall bias. However, there were only 17 cancers in women aged 65–69 years, thus this study was not powered to provide precise estimates of screening effects on cervical cancer incidence for comparisons across age groups.

The Dinkelspiel et al. (2012) study is classified as providing NHMRC level IV evidence of the effectiveness of a cytology screening exit strategy because no comparative data are presented to compare cancer incidence rates for women who had met, versus not met, exit screening criteria. Screening history was only obtained for women presenting with cancer. The investigators do compare the cervical cancer incidence rate following a negative HPV-cytology co-test with the overall incidence rate. However, there is a high risk of bias for these estimates of cancer incidence rates because the investigators relied on laboratory test use to estimate person-time at risk.

All studies were conducted in settings with organised screening programs and provide evidence applicable to Australia. Andrae et al. (2008) and Lonnberg et al. (2012 & 2013) examined the impact of five-yearly cytology screening in older women. Thus, the results have good applicability to consideration of extending the screening interval to five years for older women in Australia. Dinkelspiel et al. (2012) assessed the proportion of cervical cancers diagnosed in women over the age of 65 years who met screening exit criteria of three consecutive negative cytology tests between the ages of 55 to 65 years as is being considered for the present review. However, the cancer incidence rates in women over age 65 years reported in this health plan cohort are lower than the Australian population rates, so it is unclear how applicable the results are to the Australian setting.

Table 25 Characteristics and quality of primary studies examining screening exit age for cytology screening

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality & applicability
Andrae et al. 2008 Sweden Cancer registry and population register 1999–2001	N=7,354 • 1,230 cancers • 6,124 controls Screening • Screened and unscreened Age • 20–99 yr • Subgroups 20–53 yr 54–65 yr >65 yr	Exit age Relative effectiveness of screening vs no screening on cervical cancer incidence with comparison by age group, including >65 yr (Data for women screened <25 yr not presented)	Study design: Case control study Analysis Risk of cervical cancer • by screening participation Yes/no • and by age group Odds ratio (95% confidence interval)	ERSC Observational study NHMRC level III-2 Quality: Moderate risk of bias Screening exposure defined as screened in period up to 6 months prior to diagnosis Applicable: Fair
Lonnberg et al. 2013 Finland Cancer registry and population register 2000–2009	N=10,822 • 506 deaths • 3,036 controls Screening • Screened and unscreened Age • 25–69 yr • Subgroups 25–39 yr 40–54 yr 55–69 yr And for 5-yr age groups	Exit age Relative effectiveness of screening vs no screening on cervical cancer mortality with comparison by age group, including for age at last screen up to 65 years	Study design: Case control study Analysis Risk of cervical cancer • by screening participation Yes/no • stratified by 5-yr age group Odds ratio (95% confidence interval) Correction for self-selection bias of screening participation	Search update Observational study NHMRC level III-2 Quality: Low risk of bias Applicable: Fair
Dinkelspiel et al. 2012 United States Health insurance	N=56 cancers In women aged ≥65 yrs 1.323×10^6 women/yr	Exit age Incidence of cervical cancer for women ≥65 yr with comparison of risk by prior screening status	Study design: Health insurance plan member retrospective case series Analysis Incidence of cervical cancer for women ≥65 yr	Search update Observational study NHMRC level IV

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality & applicability
members 2003–2008	Screening • Screened and unscreened Age ≥65 yrs		Proportion of cancer cases with inadequate or adequate prior screening Where adequate prior screening = 3 consecutive negative cytology or negative co-testing between 55–65 yr	Quality: High risk of bias Applicable: Fair

Screening interval

Two primary studies investigated the relative effectiveness of screening by interval since last, or last negative, cytology test and number of prior screening tests with separate analyses performed for different age groups (Sasieni et al. 2003; IARC working group, 1986). The largest study was the UK case control study from Sasieni et al. (2003) that included 1,305 cancers diagnosed between 1990–2001 and reported results stratified in three age groups (20–39, 40–54, 55–69 years). This study was assessed as Level III-2 evidence with a low risk of bias because randomly selected controls were drawn from the same source population as cases, and all were included to avoid selection bias, with prospectively recorded screening data available for all participants to avoid recall bias.

The second study was a retrospective pooled analysis of 10 observational studies (five cohort and case control studies,) conducted in the UK, Europe and Canada that included cervical cancers diagnosed over different periods from the 1960s up to 1984, with a separate analysis for women aged under 35 years (148 cancers) (IARC working group, 1986). This study was assessed as Level III-2 evidence with a moderate risk of bias due to the potential for selection bias and recall bias in case control studies and the potential for confounding in cohort studies with historical controls that may have overestimated the risk of cervical cancer in unscreened populations.

The applicability of the Sasieni et al. (2003) study was considered good because it examined the impact of different screening intervals within a contemporary organised screening program with subgroup analyses relevant to consideration of extending screening intervals in three different age groups (although data are reported for women aged 55 years and older rather than from 50 years as specified for scenario 1). The applicability of the IARC (1986) study was considered fair because it included studies conducted when organised screening programs were first introduced in the 1960s to 1980s, and does not report separate data for older age groups of women.

Creighton et al. (2010) report on a modelled analysis to estimate the impact of adopting a three-year screening interval in Australia on cervical cancer incidence and mortality, colposcopy and biopsy rates and cost implications. This study incorporates Australian screening and treatment guidelines and integrates evidence about HPV incidence and the natural history of CIN with Australian data on cervical screening program outcomes and screening compliance. It examines the impact of screening participation using a reminder-based or call-recall system on outcomes. The results are therefore considered highly applicable to the present review question for women, but did not examine the impact of five-yearly screening in women aged 50 years and older.

Table 26 Characteristics and quality of primary studies examining screening interval for cytology screening

Author, year Setting	Population N, age	Comparison/s	Study design	Source Quality and applicability
IARC working group, 1986 International 10 studies Europe Scotland Canada 1960s–1984	Subgroup1 N=162 cancers Subgroup2 N=148 cancers Age • 35–64 yr	Screening interval Relative effectiveness of screening vs no screening with comparison by screening interval since last negative cytology and number of previous screening rounds	Study design Retrospective pooled analysis 5 case control studies , 5 cohort studies Analysis Risk of being cancer-free • by interval since last negative cytology • stratified by number of previous negative tests (1 or >1) Relative risk 95% CI ('relative protection') Subgroup analysis (i) Iceland and Aberdeen, women aged 35–64 yr (ii) Sweden, British Columbia, Manitoba, women <35 yrs, stratified by number of prior negative screens (1 vs 2+)	IARC Observational study NHMRC level III-2 Quality: Moderate risk of bias Applicable: Fair
Sasieni et al. 2003 UK Participating health authorities 1990–2001 Updates Sasieni et al. 1996	N=3,837 • 1,305 invasive cancer • 2,532 control Screening • Screened and unscreened Age • 20–69 yr • Subgroups 20–39 yr 40–54 yr 55–69 yr	Screening interval Relative effectiveness of screening vs no screening on cervical cancer incidence With comparison by screening interval since last negative cytology and by age group	Study design: Case control study Analysis Risk of cervical cancer • by screening status: Interval since last negative test vs no screening Interval since last screen vs no screening • stratified by 3 age groups Odds ratio (95% confidence interval)	IARC, AHRQ Observational study NHMRC level III-2 Quality: Low risk of bias Applicable: Good
Creighton et al. 2010 Australia	Women aged ≥65 yrs	Screening interval Analysis of the effectiveness and cost-effectiveness of extending the screening interval from 2 to 3 years	Study design: Decision model Analysis: • Difference in cervical cancer incidence, mortality; and colposcopy, biopsy and treatment rates and costs for 2-year vs 3-year screening interval • Compliance estimated using reminder-based system; or call-recall system	Search update Decision model Applicability: good Australian analysis of effectiveness and cost-effectiveness

Screening start age

HTA and systematic reviews

The two recent systematic reviews that have addressed this question did not identify definitive evidence to determine the optimal start age for population screening (Vesco et al. 2011; Peirson et al. 2012).

Vesco et al. 2011

The AHRQ review (Vesco et al. 2011) included evidence from five studies that reported on the incidence, prevalence and natural history of HPV infections and CIN in young women (Peto et al. 2004a; Woodman et al. 2001); and the benefits (Sasieni et al. 2009; Sigurdsson et al. 2007, 2010) and harms (Insinga et al. 2004) of screening in women younger than 25 years.

No studies were identified to assess risk factors for cervical cancer in young women to identify whether high risk subgroups could be defined that may benefit from an earlier screening start age. The case control study reported by Sasieni et al. (2009) was appraised as providing the best evidence about the effectiveness of screening in younger women.

The key findings of the AHRQ review were:

- In women under the age of 25 years, screening does not lead to a reduction in cervical cancer incidence diagnosed prior to 30 years, with the best available evidence indicating no screening protective effect before age 32 years.
  Based on data reported by Sasieni et al. 2009 (see Table 23).

- In women under the age of 25 years, cervical cancer is rare, and cervical screening has lower detection rates and higher false positive rates than in older women.
  Based on false positive cytology rates by age group reported by Insinga et al. 2004 (see Table 23).

For conclusions about the effectiveness of screening in women under the age of 25 years, the AHRQ review noted the small number of cancers detected by Sasieni et al. (2009) in this age group does not exclude an effect of screening to reduce the incidence of IB+ cervical cancer (for staging descriptions see Appendix D). A population-based cohort study with a historical comparison reported detection of CIN3+ in women aged 30 to 34 years declined after reducing the screening age to 20 years to provide support for a screening start age of 20 years (Sigurdsson et al. 2007, 2010), but noted limitations were the risk of confounding using a historical comparison.

To consider the possible consequences of delaying screening until 25 years, the AHRQ review also cited additional evidence about the natural history of HPV and pre-invasive cancer. They reported evidence that progression from CIN to cancer was slow from three studies that have reported progression of CIN3 to cancer at 31% in 30 years (from an unethical New Zealand study, subsequently reported by McCredie et al. (2008)), and other estimates of progression of 20–30% over 5–10 years (Chang et al. 1990; Kinlen et al. 1978). They also cited evidence about CIN regression, including a review of studies published up to 1990 that reported on regression, persistence and progression over follow-up from 1–25 years for CIN1 (57%, 32%, 1% respectively); CIN2 (43%, 35%, 5% respectively) and CIN3 (32%, 56% and >12% respectively) (Ostor 1993). A large Canadian cohort study that managed CIN2 conservatively reported 0.3% progression to invasive cancer within two years, 0.7% within five years and 1.2% within 10 years; with regression to normal cytology reported at 6.9% within two years, 29.0% within five years and

53.7% within 10 years (Holowaty et al. 1999). Rates of CIN3 progression to invasive cancer were higher (1.6% within two years, 2.6% within five years and 9.9% within 10 years (Holowaty et al. 1999)).

Based on this review, the USPSTF recommended a screening start age of 21 years based on judgment that the current evidence does not exclude potential screening benefits for women aged 20–24 years, and these potential benefits may outweigh the potential harms. This judgment was informed by the results of the Kulasingam et al. (2011) decision model that estimated the incremental colposcopies per life-year associated with screening beginning at different ages from 15 to 25 years.

Peirson et al. 2012

The Peirson et al. (2012) review included evidence from four studies that reported data on the effectiveness of screening in women in different age groups (Andrae et al. 2008; Hoffman et al. 2003; Rebolj et al. 2009; Sasieni et al. 1996, 2003, 2009). Only the Sasieni et al. (2009) case control study reported on the relative effectiveness of screening in women aged 20–24 years.

The key findings of the Peirson et al. review were:

- In women aged 30 years and older, findings of a substantial protective effect from screening were consistent across studies.
- In women below the age of 30 years, despite high participation in screening, the benefit of screening is unclear. Sasieni et al. (2009) reported no protective effect of screening for women who tested between ages 20 and 21 or ages 22 to 24 years.
- In women below the age of 30 years, 10% have abnormal results resulting in further testing and associated minor harms; 1.5% are referred to colposcopy, and 50% of these women are referred to treatment with the potential for more serious harms such as future early pregnancy loss or premature labour.

The authors noted decisions about the start age for screening in women younger than 30 years requires judgments about an acceptable trade-off between potential benefits versus harms.

Based on this review, the CTFPHC recommended a screening start age of 25 years due to the lack of evidence of the effectiveness of screening in women younger than 25 years, compared to moderate quality evidence that cytology screening may have a small effect in reducing cervical cancer morbidity and mortality in women 25–29 years. They noted screening decisions for young women must consider the balance between potential benefits and potential harms.

Additional evidence used to make this judgment included:

- Cancer registry data showed no reduction in cervical cancer mortality in women aged 20–24 years since screening was introduced in the 1970s.
- Women younger than 30 years are at higher risk of undergoing additional diagnostic procedures and potentially unnecessary treatment because the specificity of cytology for detecting CIN is lower and the risk of false positive tests is higher in this age group.

The CTFPHC assessed their recommendation to commence screening at 25 years as weak given the uncertainty of supportive evidence.

Australian data

The NCSP has reported reductions in age-specific cervical cancer incidence and mortality rates across all age groups since the program commenced in 1991 (AIHW 2012). As shown in Figure 4 and Figure 5, page 11, this impact has been greatest in older women. For women aged 20–24 years, cervical cancer is very rare and the magnitude of difference in incidence in absolute terms is very small, with cancer registry data showing a five-year age specific cervical cancer incidence in the pre-screening period 1983–1987 of 2.7 per 100,000 compared to 1.4 per 100,000 in 2002–2007. Both screening and changes in population risk factors may have contributed to this small difference in cancer incidence, as demonstrated in the Patel et al. 2012 control group (as outlined below and in Table 27); therefore, it is not possible to use these data to estimate the impact of raising the screening age to 25 years. Even so, these data indicate any benefits of the screening program for reducing cancer incidence in this age group are not larger than ~1 cancer avoided per 100,000 women.

Importantly, over the same 25-year period, cervical cancer mortality rates in women aged 20–24 years have not changed with rates of between 0–0.2 deaths per 100,000 women reported in 5-year periods before and after the screening program was introduced (Figure 5). Over the same period, reductions in both cervical cancer incidence and mortality have been observed in women aged 25–29 years. As outlined below, Sasieni et al. (2009) have investigated the evidence whether screening women aged 20–24 years has contributed to reducing cervical cancer incidence in women aged 25–29 years.

Primary studies

Results from the five primary studies examining the impact of changing the cytology screening start age from 20 to 25 years are summarised in Table 27 to Table 32.

Discussion

The UK case control study reported by Sasieni et al. (2009) provides the highest quality evidence available for a comparison of the effectiveness of screening in women 20–24 years versus 25–29 years. This study observed screening women aged 22–24 years does not lead to a reduction in the incidence of cervical cancer between ages 26.5–30 years. In addition, the risk of cervical cancer was not lower for women with two screening episodes before the age of 25 years versus no screening. In contrast, screening older age groups (from age 32 years) was associated with a substantial reduction in cervical cancer incidence (Table 27). This study provides strong evidence that screening is less effective in women under the age of 25 years than older age groups.

Table 27 Relative effectiveness of screening versus no screening for women by age group (Sasieni et al. 2009)

Age of screening vs no screening	Age of cancer diagnosis	Odds ratio (95% CI)	Interpretation
Any screen 22–24 yr vs nil screen 20–24 yr	25–29 yr	1.11 (0.83 to 1.50)	Screening women 22–24 yr does not reduce incidence of a cervical cancer diagnosis <30 yr
Two screens 20–22 yr & 23–25 yr vs 1st screen 23–25 yr	26.5–29 yr	1.1 (0.62 to 2.0)	2 vs 1 screening episodes does not reduce cancer diagnosis in age 26.5–29
Any screen 32–34 yr vs nil screen 30–34 yr	35–39 yr	0.55 (0.44 to 0.69)	The effectiveness of screening for women<32 years is not demonstrated.
Any screen 42–44 yr vs nil screen 40–44 yr	45–49 yr	0.37 (0.29 to 0.48)	Screening older age groups provides substantial risk reduction for women screened in previous 5–8 years
Any screen 52–54 yr vs nil screen 50–54 yr	55–59 yr	0.26 (0.19 to 0.36)	-

However, despite its large size, due to the very small numbers of cancers in women under the age of 25 years, this study does not definitively exclude the possibility of a small screening benefit in women under the age of 25 years. In particular, study data do not exclude the possibility that screening will shift diagnosis to early microinvasive disease, reducing the (small) risk of being diagnosed with invasive IB+ cancer in subsequent years.

Two UK studies of temporal trends in cervical cancer incidence before and after the screening start age was raised to 25 years in England provide evidence that raising the screening start age to 25 years does not lead to an increase in cervical cancer incidence when compared to Wales and Scotland where the screening start age remains at 20 years (Sasieni & Castanon 2012; Patel et al. 2012). As expected, both studies report the incidence of high-grade CIN decreased in women aged 20–24 years after the screening age was raised in 2004, with an associated increase in the incidence of high-grade CIN in women aged 25–29 years. Both studies also documented an increase in cervical cancer incidence after 2004. However, these cancer incidence trends were also observed in the comparator countries that continued to start screening at age 20 years, indicating the impact of factors other than screening age. In addition, analysis of trends of cervical cancer incidence in England by birth cohort indicate that the increase in incidence observed in women in this age group precedes the change in screening start age (Foley et al. 2011; see 'Age-stratified incidence and mortality' under 'Cervical cancer trends in Australia', page 10).

Patel et al. (2012) reported an increase in cervical cancer incidence in young women in both age groups (20–24 years and 25–29 years) relative to women aged >30 years 2005–2009 versus 2000–2004. However, they also documented a similar trend over time in young women in Wales where the screening was unchanged (Table 28). Furthermore, the investigators presented data on the increasing incidence of STD over the same time period in both populations showing higher rates in the English population, indicating the increase in cervical cancer incidence over time may be attributed to changes in risk factors.

Table 28 Impact of extending screening start age from 20 to 25 years in 2004 on cervical cancer incidence, 2000–2009 (Patel et al. 2012)

Region	Age group	HG-CIN Incidence rate /10^5 Pre vs post 2004	HG-CIN IRR (95% CI) by time period and age group	Cervical cancer Incidence rate /10^5 Pre vs post 2004	Cervical cancer IRR (95% CI) by time period	Cervical cancer IRR (95% CI) by time period and age group	Cervical cancer Time period, age group interaction
North-east England	20–24 yr	432.2 vs 230.7	0.56 (0.51–0.61)	3.4 vs 7.0	-	2.18 (1.09–4.37)	0.028
-	25–29 yr	472.0 vs 603.0	1.34 (1.24–1.45)	12.0 vs 21.0	1.21 (1.08–1.34)	1.93 (1.28–2.92)	0.002
-	30–34 yr	64.9 vs 61.5	1111111.00	18.4 vs 25.5	-	1.51 (1.08–2.13)	0.017
-	35+ yr	-	-	10.6 vs 9.8	-	1.00	Reference
Wales (contemporary control group)	20–24 yr	219.0 vs 310.9	1.21 (1.10–1.33)	2.3 vs 5.9	-	2.72 (1.31–5.62)	0.007
-	25–29 yr	298.7 vs 431.0	1.23 (1.13–1.35)	11.5 vs 19.8	1.17 (1.06–1.30)	1.82 (1.26–2.65)	0.002
-	30–34 yr	52.1 vs 60.9	1.00	16.0 vs 20.1	-	1.33 (0.96–1.84)	0.086
-	35+ yr	-	-	15.6 vs 14.8	-	1.00	Reference

Sasieni and Castanon (2012) included a larger sample of cervical cancers reported across England between 2000 and 2010 and compared incidence rates with both Wales and Scotland (and included data to 2011 for comparisons of time trends by year of age in England (Table 30)). They reported no difference in the trend of increasing cervical cancer incidence in women aged 25–29 years in England versus Wales and Scotland over this period (RR 0.98, 95% CI 0.69–1.39) (Table 29).

Table 29 Impact of extending screening start age from 20 to 25 years in 2004 on cervical cancer incidence: trends per decade in cervical cancer incidence by age group and nation: 2000–2010 (Sasieni & Castanon 2012)

Age group	England (RR (95% CI)	Scotland (RR (95% CI)	Wales (RR (95% CI)	England vs Wales & Scotland	P-value
20–24 yr	0.98 (0.75–1.29)	1.39 (0.64–3.05)	2.16 (0.82–5.66)	0.59 (0.30–1.15)	0.12
25–29 yr	2.37 (2.08–2.70)	2.96 (1.94–4.53)	1.77 (1.05–2.99)	0.98 (0.69–1.39)	0.90
30–34 yr	1.49 (1.34–1.66)	2.67 (1.90–3.75)	1.29 (0.81–2.05)	0.72 (0.54–0.96)	0.03
35–39 yr	1.24 (1.12–1.38)	2.47 (1.77–3.45)	1.15 (0.77–1.73)	0.68 (0.51–0.90)	0.01

Table 30 Impact of extending screening start age from 20 to 25 years in 2004 on cervical cancer incidence: number of cancers diagnosed in each financial year by age at diagnosis, England, N=4,079 cancers (Sasieni & Castanon 2012)

Age	2007/2008	2008/2009	2009/2010	2010/2011	Trend (RR/yr)	P value
20–22 yr	14	10	7	5	0.71	0.028
23 yr	15	6	13	11	0.96	0.74
24 yr	23	15	18	17	0.92	0.43
25 yr	36	42	91	145	1.67	<0.0001
26 yr	45	50	57	56	1.08	0.22
27–29 yr	147	196	236	163	1.05	0.15
30–34 yr	298	322	342	306	1.01	0.58
35+ yr	1552	1563	1585	1585	1.01	0.5

In addition, a population-based registry cohort study examined the impact of reducing the start age in Iceland from 25 to 20 years (Sigurdsson 2007, 2010). Overall, the investigators did not find evidence of a reduction in cervical cancer incidence over time in women in 20–24, 25–29 or 30–34-year age groups to demonstrate the effectiveness of lowering the screening start age to 20 years. However, they reported an increased incidence of stage IA cancers in women aged 20–34 years, and reduced incidence stage IIA suggesting a shift to earlier disease detection potentially attributable to earlier screening (Table 31).

Table 31 Cervical cancer in Iceland 1964–2008: age-specific incidence rates at ages 20–29 and 20–34 per 100 000 before and after changing the lower age limit from age 25 to 20 in 1988. (Sigurdsson 2010)

Age group	All cancers Pre vs post 1988	Stage IA Pre vs post 1988	Stage IB Pre vs post 1988	Stage IIA+ Pre vs post 1988	Screen detected Pre vs post 1988	Clinically detected Pre vs post 1988
20–24 yr	2.1 vs 2.8 *p* =0.6	-	-	-	-	-
20–29 yr	6.7 vs 9.7 *p* =0.1	2.7 vs 6.6 *p* =0.008	2.7 vs 2.8 *p* =0.9	1.1 vs 0.2 *p* =0.1	3.0 vs 7.6 *p* =0.003	3.7 vs 2.1 *p* =0.2
20–34 yr	10.7 vs 13.0 *p* =0.2	4.7 vs 8.9 *p* =0.005	3.6 vs 4.0 *p* =0.7	2.4 vs0.2 *p* =0.001	5.2 vs 10.5 *p* <0.001	5.5 vs 2.5 *p* =0.009

Screen-detected cases are stage IA and subclinical (occult) IB cases; clinical cases are all other cases. P-values for rate differences in 1964–1988 versus 1989–2008. Incidence rate/10^5

p-value for test of difference in incidence rate by time period

Evidence about the relative harms for screening in young versus older women is more definitive. Insinga et al. (2004) demonstrated the risk of false positive cytology results is higher for women aged 20–24 years than women aged 25–29 years and older age groups (Table 32). These results reflect the high prevalence of HPV infection in this age group (see 'Natural history of cervical cancer', page 32) and the lower specificity of cytology in women under 30 years (see Table 87). In addition to the inconvenience of further testing for abnormal cytology results where no abnormalities are found on follow-up, in cases where CIN changes lead to treatment, treatment harms in young women, in particular pregnancy adverse outcomes, for lesions that may otherwise regress also need to be taken into account (see 'Harms of screening', page 23).

Table 32 Abnormal cytology tests as a proportion of all routine cytology tests (Insinga et al. 2004)

Age group	N Cytology tests	FP%	CIN1%	CIN2%	CIN3%
20–24	852	3.5	0.9	0.6	0.2
25–29	1,952	2.1	0.2	0.6	0.6
30–39	5,992	2.6	0.5	0.3	0.4
-	-	1.6	0.1	0.0	0.0
60–69	3,543	1.8	0.1	0.0	0.1
70–79	1,657	-	-	-	-

Note: Of 30,936 routine cytology tests, 1,331 were abnormal tests: 51.5% false positive, 29% incomplete follow-up, 19% CIN/cancer

In Australia, NCSP data also show that among women aged 20–69 years, women aged 20–24 years have the highest risk of abnormal cytology results, the second highest risk of high-grade histology after women aged 25–29 years, but the lowest risk of cervical cancer (see Table 8and Table 9, Figure 4 and Figure 5, pages 33 and 11; AIHW, 2013). These data indicate that screening women aged 20–24 years is more likely to lead to further investigation and potentially treatment than older age groups, despite their low risk of cancer.

In the absence of direct evidence about the optimal age for starting screening, recommendations require judgments about the age at which screening benefits outweigh the potential for harms. Both the USPSTF and the CTFPHC concluded against screening in women aged 20 and younger based on evidence that this population are at very low risk of cervical cancer, and the harms of tests, procedures and in some cases treatment for potentially transient changes were judged to outweigh the benefits. However, they differed on recommendations for screening women aged 20–24 years due to ongoing uncertainty about the possibility of modest screening benefits in this age group, and uncertainty about the acceptable trade-off between the potential benefits of early detection and treatment versus the potential harms.

In Australia with the introduction of the HPV vaccination program, lower rates of HPV and pre-invasive cancers are anticipated. Early data from the Victorian Cervical Cytology Registry have shown a decrease in the number of histologically confirmed high-grade cervical lesions detected in women aged 20–24 from 19.8 per 1,000 in 2008 to 15.7 per 1,000 in 2011 (VCCR 2011). The need to balance uncertain modest screening benefits against the risk of harm, and explore the potential impact of vaccination, highlights the value of undertaking an Australian decision model analysis to synthesise the evidence available.

Screening exit age

HTA and systematic reviews

The two recent systematic reviews did not identify direct evidence to determine the optimal exit age for population screening (Peirson et al. 2012; Saslow et al. 2012).

Peirson et al. 2012

The Peirson et al. (2012) review identified no studies designed to provide evidence about the optimal age to exit screening. The reviewers identified one study that reported data on the relative effectiveness of screening in women older than 65 years (Andrae et al. 2008). The reviewers also reported supportive data that screening exposure rates were lower among cases aged 65–74 years than controls, but similar for women aged 75 years and older (Sasieni et al. 1996).

The key findings of the Evidence Review and Synthesis Centre (ERSC) review were:

- In women over the age of 65 years, evidence from one study indicates exposure to cytology screening continues to provide a strong protective effect.

Based on this review, the CTFPHC recommended a screening exit age of 70 years for women who have had three consecutive negative cytology tests in the past 10 years; and ongoing screening for older women until three consecutive negative tests have been reported. They noted that the lack of evidence available about the effectiveness of screening in women aged 70 years and older reflected the fact that most studies excluded women in this age group. In making their recommendation for requiring ongoing screening for women with an inadequate prior history of screening, they noted that cervical cancer incidence and mortality remained high in older age groups.

Additional evidence used to make the judgment about screening to 70 years of age was the evidence of less harm from screening for older women.

The CTFPHC assessed their recommendation to exit screening at 70 years as weak given the uncertainty of supportive evidence.

Saslow et al. 2012

The Saslow et al. (2012) review included evidence from one modelling study that estimated the effectiveness of no further screening after adequate negative screening to age 65 years versus screening with cytology every three years or every five years (Kulasingam et al. 2011).

The key findings of the Saslow et al. (2012) review based on this study were:

- For women aged 65 years who have been screened every three years, continued screening even to age 90 prevents 1.6 cancers and 0.5 cancer deaths per 1,000 women.
- Screening after age 65 years was estimated to extend life expectancy by one year per 1,000 women (less than one day per woman) while resulting in 58 extra false positives that require investigation, 127 extra colposcopies and 13 extra CIN2/3 requiring treatment compared with a strategy of exiting screening aged 65 years.

In addition to this evidence, the reviewers also noted:

- A previous modelling study found that estimating screening women after the age of 65 years is inefficient for those who have an adequate screening history, but reduces mortality for women who have not been adequately screened (Fahs et al. 1992).
- Data that most cervical cancers in the US in women over the age of 65 years are diagnosed in unscreened or infrequently screened women (Mandelblatt et al. 1986; Sawaya et al. 2000).

- Consideration that screening older women may be less effective than younger women because the transformation zone of the cervix in older women is smaller and less accessible for screening; and it is less likely for new cases to have time to progress to invasive cancer in the woman's lifetime.
- Evidence that cytology sampling is associated with potential harms, including anxiety and discomfort which may be greater in some older women due to vaginal atrophy and cervical stenosis.

Based on this evidence, the ACS-ASCCP-ASCP recommended a screening exit age of 65 years for women who have three consecutive negative cytology tests or two consecutive negative HPV/cytology co-tests in the last 10 years if the most recent test occurred in the past five years; and no history of CIN2+ or cervical cancer within the last 20 years.

They assessed their recommendation as weak, because the choice of exit age was based on the judgment of the expert panel members about a favourable balance between the benefits and harms of screening older women.

Australian data

The NCSP has reported reductions in age-specific cervical cancer incidence and mortality rates across all age groups since the program commenced in 1991 (AIHW 2012). As shown in Figure 4 and Figure 5, page 11, this impact has been greatest in older women. However, it is not possible to estimate to what extent screening women aged 65–69 years has contributed to the improvements observed in women aged over 65 years. The evidence identified to address this question is discussed below.

NCSP data also show that approximately 50% of women aged 65–69 years participate in the cervical program compared to 63% of women aged 45–49 years. Women aged 65–69 year who present for screening have a low (1%) risk of abnormal cytology results, but account for 7% of cancers detected, and overall women in this age group continue to have a moderate risk of cervical cancer (11.4 per 100,000).

Data from the Victorian Cervical Cytology Register (VCCR) indicate that 80% of Victorian women diagnosed with invasive cervical cancer in 2009 were underscreened (VCCR 2011). However, these data are not reported stratified by age.

Primary studies

Results from the three primary studies examining the effectiveness of screening women after the age of 65 years, and the impact of exiting screening at 65 years are summarised in Table 33 to Table 35.

Discussion

Overall, no studies were identified that provided definitive evidence to compare the relative effectiveness of exiting screening at 65 years versus 70 years. The data available provide evidence that screening at age 65 years is associated with a reduced incidence of cervical cancer incidence and mortality.

The Swedish case control study reported by Andrae et al. (2008) provides moderate quality evidence and the largest sample size to examine the effectiveness of screening in women with a cancer diagnosis after the age of 65 years. However, the exclusion of screening undertaken within the six months prior to the cancer diagnosis may overestimate the effect of screening. The investigators observed a statistically significant protective effect from screening in this age group (odds ratio (OR)

2.79 95% CI 1.89–4.11 for the risk of cervical cancer in women with no screening 6 to 66 months prior to diagnosis versus screening in this period).

Table 33 Odds of cervical cancer in women not screened versus screened within recommended interval (Andrae et al. 2008)

Age at diagnosis	Screening exposure within recommended interval (yes/no)	Screening exposure within recommended interval (yes/no)	Odds ratio (OR) (95% CI) Not screened vs screened
-	Cases	Controls	-
All ages	441/789	3,288/2,836	2.52 (2.19–2.91)
21–29	26/37	189/120	2.37 (1.36–4.13)
30–65	383/394	2,733/1,142	2.51 (2.14–2.94)
>65	32/358	366/1,574	2.79 (1.89–4.11)
-	-	-	P value heterogeneity = 0.96

These findings were supported by a high-quality smaller case control study that reported a similar risk reduction for screening in reducing the risk of a cancer diagnosis in women screened aged 65–69 years (OR 0.49, 95% CI 0.10–2.41, corrected for self-selection bias) and women screened aged 60–64 years (OR 0.49, 95% CI 0.28–0.84, corrected for self-selection bias), although this effect was not statistically significant for older women with only 17 cancer cases identified (Lonnberg et al. 2012). Data to examine the effectiveness of screening after age 65 years in women who have had three prior negative screens prior to 65 years were not reported.

Table 34 Odds ratios of the association between cervical cancer incidence and screening participation (Lonnberg et al. 2012)

Age at invited screening	Screening exposure: yes/no	Screening exposure: yes/no	Odds ratio (OR) (95% confidence interval) for cervical cancer incidence	Odds ratio (OR) (95% confidence interval) for cervical cancer incidence
-	Cases	Controls	Crude OR	OR corrected for self-selection bias
Overall	366/494	3,469/1,813	0.37	0.53 (0.46–0.62)
25–29	11/18	77/77	0.65	0.95 (0.36–2.49)
30–34	50/75	422/351	0.54	0.79 (0.53–1.16)
35–39	61/80	529/377	0.50	0.73 (0.50–1.06)
40–44	44/87	527/296	0.26	0.38 (0.26–0.58)
45–49	50/66	514/216	0.28	0.41 (0.27–0.62)
50–54	53/57	486/183	0.30	0.44 (0.29–0.68)
55–59	49/63	497/165	0.23	0.34 (0.22–0.52)
60–64	40/39	350/128	0.33	0.49 (0.28–0.84)
65–69	8/9	67/20	0.33	0.49 (0.10–2.41)

An analysis of the impact of screening on cervical cancer mortality over the same period demonstrated a strong protective effect for screening at age 65 years in women with a diagnosis of cervical cancer age 65–80 year (Lonnberg et al. 2013).

Table 35 Odds ratios of the association between cervical cancer mortality and screening participation (Lonnberg et al. 2013)

Age at last invited screening for cancers diagnosed >65 years	Cases Screened: yes/no	Controls Screened: yes/no	Crude Odds ratio (OR)	Odds ratio (OR) (95% CI) for cervical cancer mortality corrected for self-selection bias
55 yr	2/2	17/14	1.22	2.45 (0.15–39.7)
60 yr	14/16	138/42	0.27	0.54 (0.21–1.34)
65 yr	7/5	74/23	0.14	0.28 (0.03–2.47)

-	Age at last screen	Cases	Controls	Odds ratio (95% CI)
For diagnosis age 66–80 yr	None (no screen ≥55)	156	840	reference
-	55 yr	2	17	0.45 (0.09–2.17)
-	60 yr	14	138	0.41 (0.21–0.78)
-	65 yr	6	73	0.38 (0.16–0.90)
For diagnosis age 71–80 yr	None (no screen ≥55)	127	712	reference
-	55 yr	0	0	N/A
-	60 yr	5	50	0.43 (0.15–1.23)
-	65 yr	2	42	0.24 (0.06–1.00)

A retrospective analysis of 56 cervical cancers in members of a US health fund provided evidence that cervical cancers are diagnosed in women aged over 65 years who met the exit criteria of having three consecutive negative cytology tests in the previous 10 years (14 cases, 25%) but these are the minority, with 75% of cervical cancers detected in women with inadequate prior screening histories (Table 36, Dinkelspiel et al. 2012).

Table 36 Invasive cervical cancers in women age 65 and older by five-year age interval and incidence (Dinkelspiel et al. 2012)

Age at cancer diagnosis	N Cancer	Incidence rate/10^5 women-year	Met exit criteria# Yes N (%)	Met exit criteria# No N (%)
All ages ≥65 yrs	56	4.2	14 (25%)*	42 (75%)
≥65 yrs + negative co-test	3	2.3	-	-
65–69 yrs	18	4.7	-	-
70–74 yrs	14	4.4	-	-
75–79 yrs	8	3.0	-	-
80–84 yrs	12	6.0	-	-
85+ yrs	4	2.5	-	-

exit criteria at 65 years = 3 consecutive negative cytology tests aged 55–65 years, or single negative HPV & cytology co-test
*All 14 had at least 3 negative cytology tests prior to diagnosis; 2 had last cytology age 62 years, 12 had last cytology aged >65 years

Screening interval

HTA and systematic reviews

The two recent systematic reviews identified evidence about the impact on cervical cancer incidence of extending the screening interval for different age groups (Peirson et al. 2012; Saslow et al. 2012).

Neither review identified direct evidence to compare the effectiveness of extending the screening interval from two to three years for women under the age of 50 years, and to five years for women aged 50 years and older.

Peirson et al. 2012

The Peirson et al. (2012) review identified 14 studies reporting on the effectiveness of screening on cervical cancer incidence at different screening intervals ranging from six months to over 10 years, of which three studies reported data for different age groups (Andrae et al. 2008; Sasieni et al. 2003; Rebolj et al. 2009).

The key findings based on this evidence were:

- In women aged 30 years and older, the protective effect of screening is substantial and consistently observed for screening programs using intervals of up to five years.
- Cytology screening using shorter intervals offered greater protection than longer intervals, but intervals of 10–15 years were still associated with a lower risk of cervical cancer than no screening.

Based on this review, the CTFPHC recommended a 3-yearly screening interval for women aged 25–69 years.

The CTFPH assessed their recommendation as strong for women aged 30–69 years (and weak for women under the age of 30 years because the effectiveness of screening has not been strongly established in this age group). They noted this recommendation was based on the task force's judgment that the small incremental benefit of three-yearly screening versus a longer interval balances the greater potential for harm from increased testing and procedures.

Two studies reported data relevant to the age groups specified for the present review and are further summarised below (Andrae et al. 2008; Sasieni et al. 2003).

Saslow et al. 2012

The Saslow et al. (2012) review included evidence from 13 primary studies and five modelling studies.

The key findings were:

- For women aged 21–29 years, modelled estimates indicated screening every three years is associated with a lifetime cervical cancer risk to 5–8 cancers per 1,000 women, compared to 4–6 incidence cancers per 1,000 women for screening every two years. The lifetime risk of death due to cervical cancer is estimated at 0.05 and 0.05 deaths per 1,000 women using three- and two-year intervals respectively.
- For women under age 30 years, screening every three years is predicted to lead to 760 colposcopies per 1,000 women, compared to approximately 1,080 colposcopies per 1,000 women if two-year screening.
- For women aged 30–65 years, screening intervals of greater than three years were associated with an increased risk of cervical cancer.
- Two primary studies did not find that the number of prior negative cytology tests modified the association between screening interval and risk of cancer (Gram et al. 1998; Miller et al. 2003).

Based on this evidence, the ACS-ASCCP-ASCP recommended cytology screening every three years for women aged 21–29 years. For women aged 30–65 years, cytology screening every three years was recommended as acceptable, but co-testing with cytology and HPV every five years was recommended as preferable. The working group found insufficient evidence to recommend longer cytology screening intervals for women who have two previous consecutive negative tests.

Australian data

An Australian case control study reported on the screening behaviour of 877 women aged 20 to 69 diagnosed with invasive cervical cancer in NSW between 2000 and 2003 (Yang et al. 2008). This study compared the risk of developing invasive cervical cancer in women who had one Pap test in the preceding four years ('irregular' screening) and Pap tests in two or more years of the preceding four years ('regular' screening) to that in women who had not participated in Pap testing during this period. The authors used logistic regression to adjust for other potential confounders including Pap test result in the preceding six years. The relative risk (RR) for invasive cervical cancer was significantly reduced by both irregular and regular screening, with regular screening providing a greater benefit (irregular screening RR 0.19, 95% CI 0.13–0.27; regular screening RR = 0.07, 95% CI 0.04–0.10). However, these data do not provide a clear comparison of two versus three-yearly screening. The evidence identified to address this question is discussed below.

Primary studies

Results from the three primary studies examining the impact of screening every three years versus two years for women under the age of 50 years and every five years versus two to three years are summarised in Table 37.

Discussion

Overall, the IARC working group (1986) pooled analysis showed that the protective benefit of screening is stronger when using shorter screening intervals (Table 36,Table 37 and Table 39). For women under 35 years of age, including the subset who had two or more prior negative screening tests, the risk reduction offered by screening was diminished substantially for intervals greater than three years (Table 37).

Table 37 Relative protection against cervical cancer in women under age 35 in Sweden and Canada (IARC working group, 1986)

Time since last screen, years -	N Cancers[1] -	Relative risk cancer-free (95% CI not reported) 1 prior negative test	Relative risk cancer-free (95% CI not reported) **≥2 prior negative test**
0–0.9	9	11.7	16.8
1–1.9	18	4.9	10.9
2–2.9	16	4.9	10.6
3–3.9	29	2.5	3.8
4–4.9	18	2.7	2.4
5–5.9	16	1.7	-
6–6.9	9	2.1	-
7–7.9	10	1.2	-
≥8	5	1.5	-
Never screened	-	1.00	-

1 For women <35 yr, most cases 30–34 yr. Assuming incidence of $20/10^5$ in historical unscreened population

Table 38 Estimated impact of screening program using different age ranges and intervals, based on full compliance—data from Scotland and Iceland, women age 35–64 years (IARC working group, 1986)

Time since last screen, years	N Cancers	Relative risk cancer-free (95% CI)	Impact of screening interval on % reduction in cumulative incidence of cervical cancer and no. of tests performed Given 2nd negative cytology at 35 years
0–0.9	25	15.3 (10.0–22.6)	1-year interval = 93.5% (30 tests)
1–1.9	23	11.9 (7.5–18.3)	2-year interval = 92.5% (15 tests)
2–2.9	25	8.0 (5.2–11.8)	3-year interval = 90.8% (10 tests)
3–3.9	30	5.3 (3.6–7.6)	-
4–4.9	30	2.8 (1.9–4.0)	5-year interval = 83.6% (6 tests)
5–5.9	16	3.6 (2.1–5.9)	-
6–6.9	6	1.6 (0.6–3.5)	-
10	7	0.8 (0.3–1.6)	10-year interval = 64.1% (3 tests)
Never screened	-	1.00	-

Table 39 Effect of different screening policies on incidence rates of cervical cancer in women aged 20–64 (IARC working group, 1986)

Age	Screening Interval	Cumulative incidence rate/10^5 women	% reduction in incidence	No. of tests	No. of cases prevented/10^5 tests
20–64	No screening	1575.0	–	–	–
20–64	Every year	105.2	93	45	33
20–64	Every 3 years	138.9	91	15	96
25–64	Every 3 years	161.8	90	13	109
35–64	Every 3 years	354.9	78	10	122
20–64	Every 5 years	258.6	84	9	146
25–64	Every 5 years	287.8	82	8	161
35–64	Every 5 years	480.9	70	6	182

Sasieni et al. (2003) reported more applicable clinical data for the age groups specified for the present review. They found a screening interval of up to three years was effective for reducing cervical cancer incidence in women below the age of 40 years (OR 0.28, 95% CI 0.20–0.41), whereas a screening interval of three to five years was not associated with a reduction in cervical cancer incidence in this age group (Table 40). In contrast, a screening interval of three to five years was effective for older women (women aged 40–54 years OR 0.39, 95% CI 0.26–0.58; women 55–69 years OR 0.20, 95% CI 0.12–0.33). Screening was also shown to be effective for intervals longer than five years for women aged 55–69 years. These results were used to estimate that in women aged 20–39 years, 61% of cancers in women would be prevented by three-yearly screening, compared to 30% by five-yearly screening (Table 41). The additional protective benefits of a shorter screening interval diminished for older women, so that for women aged 55–69 years, 87% versus 83% of cancers were prevented by three and five-yearly screening respectively.

Table 40 **Odds ratios for invasive cervical cancer by time since the last negative cytological smear and age (Sasieni et al. 2003)**

Years since last negative cytology	Age at cancer diagnosis: 20–39 years Odds ratio (95% CI)	Age at cancer diagnosis: 40–54 years Odds ratio (95% CI)	Age at cancer diagnosis: 55–69 years Odds ratio (95% CI)
No prior negative cytology	1.00	1.00	1.00
0–1.5	0.24 (0.16– 0.37)	0.12 (0.08–0.18)	0.13 (0.08–0.22)
1.5–2.5	0.33 (0.21–0.51)	0.14 (0.08–0.22)	0.13 (0.07–0.23)
2.5–3.5	0.67 (0.43 –1.04)	0.25 (0.16–0.40)	0.15 (0.08–0.26)
3.5–4.5	1.06 (0.65–1.72)	0.30 (0.18–0.50)	0.18 (0.09–0.34)
4.5–5.5	1.40 (0.75–2.62)	0.61 (0.34–1.09)	0.28 (0.14–0.57)
5.5–6.5	1.86 (0.88–3.93)	0.72 (0.36–1.43)	0.33 (0.14–0.79)
Over 6.5	2.37 (1.16–4.85)	0.69 (0.36–1.34)	0.55 (0.27–1.10)
One negative only	1.00 (0.70–1.41)	1.31 (0.90–1.91)	1.39 (0.88–2.21)
<3 yr	0.28 (0.20–0.41)	0.12 (0.08–0.17)	0.13 (0.08–0.19)
3–5 yr	1.03 (0.68–1.56)	0.39 (0.26–0.58)	0.20 (0.12–0.33)
>5yr	2.05 (1.20–3.49)	0.72 (0.43–1.18)	0.45 (0.25–0.81)
One negative only	1.03 (0.73–1.46)	1.32 (0.91–1.91)	1.35 (0.85–2.14)

Table 41 **Summary percentage of preventable cervical cancer by three- and five-yearly screening and age based on Table 40 (Sasieni et al. 2003)**

Screening	20–39 yr	40–54 yr	55–69 yr
Annual screening	76%	88%	87%
3-yearly screening	61%	84%	87%
5-yearly screening	30%	73%	83%

The Australian modelling study, Creighton et al. (2010), reported on the impact of extending the screening interval from two to three-yearly screening taking into account current screening participation rates. This analysis predicts that extending the screening interval to three years will not substantially alter cervical cancer incidence or mortality rates but will lead to a reduction in the number of cytology tests and colposcopy procedures (Table 42). These results are consistent with results from a US-modelled analysis (Kulasingam et al. 2011).

Table 42 **Observed and predicted outcomes from two versus three-yearly screening intervals in Australia (Creighton et al. 2010)**

Screening outcomes	% difference 3 vs 2 years	N 3 vs 2 years
Av. lifetime risk of cervical cancer	~ 0.68% difference	-
Annual incidence	-	731–2 vs 736
Annual mortality	-	204–5 vs 208
Cytology tests	7–13% decrease	140,000–250,000
Colposcopy	4–10% decrease	2,700–6,400 fewer
Biopsies	-	1,400–3,200 fewer
Number of treatments for CIN2/3	2–4% decrease	300–600 fewer

Results of Assessment— Effectiveness: Liquid-based cytology (scenario 2)

Manual LBC for primary screening

Primary research question 2

What is the comparative safety, effectiveness and cost-effectiveness of either filtration or cell enrichment LBC (using the IARC recommendations for age range and interval for cytology), compared with the protocol used in the current Australian cervical screening program?

Summary 5: Summary of effectiveness – Liquid based cytology

No studies have assessed the impact of LBC with manual reading on incidence or mortality rates of invasive cervical cancer compared to conventional cytology. Therefore, a modelled analysis of cervical cancer screening, diagnosis and treatment using the IARC recommendations in HPV-vaccinated and unvaccinated populations is in progress to explore the potential long-term benefits and harms of this technology in the Australian setting.

Test accuracy

MSAC (2009) concluded that LBC provides no statistically significant difference in sensitivity (at HSIL, LSIL or pLSIL thresholds) or specificity (at HSIL or LSIL thresholds) for the detection of CIN2+ compared to conventional cytology, but that LBC reduces the specificity for the detection of CIN2+ at a test threshold of pLSIL. This conclusion was based upon a high-quality systematic review by Arbyn et al. (2008).

The relative accuracy data from an applicable RCT published since the 2009 review (the NETHCON trial) are in accord with these conclusions.

Unsatisfactory rates and test yields

The MSAC (2009) report concluded that LBC reduces that rate of unsatisfactory smears in comparison with conventional cytology and that LBC classifies significantly more slides as positive for low-grade lesions. This conclusion was based upon a HTA and meta-analysis by Krahn et al. (2008). The data presented in the current review from an applicable RCT published since the 2009 review (the NETHCON trial) are in accord with these conclusions.

Conclusions

- LBC provides no statistically significant difference in sensitivity (at HSIL, LSIL or pLSIL thresholds) or specificity (at HSIL or LSIL thresholds) when compared with conventional cytology (MSAC 2009).
- LBC reduces the rate of unsatisfactory smears in comparison with conventional cytology (MSAC 2009).

Table 43 Inclusion criteria for identification of studies relevant to assessment of effectiveness of LBC

Selection criteria	Inclusion criteria
Population	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Intervention	Manual screening of LBC
Comparator(s)	Conventional cytology
Outcomes	Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC) in women with a possible or definite HSIL cytology result Cervical cancer incidence Mortality
Reference standard	Colposcopy with biopsy for positives
Search period	2008–2012 (HTAs); 2011–2012 (primary studies)
Language	English
Study design [a]	RCT, pseudorandomised trial

Further detail provided in Box 2, page 39.

[a] Post-hoc inclusion criteria based upon the body of evidence identified.

Abbreviations: AIS = adenocarcinoma in situ; HSIL = high-grade squamous intraepithelial lesion; HTA = health technology assessment; LBC = liquid-based cytology; SCC = squamous cell carcinoma

The most recent systematic reviews and HTAs identified and included for the assessment of evidence for LBC are listed in Table 44. The inclusion criteria used are listed in Table 43.

Table 44 Relevant systematic reviews and meta-analyses included for effectiveness of LBC versus CC for primary cervical cancer screening

Study identified in review	Study title	Studies included in this review
Peirson et al. 2012	Peirson, L., Fitzpatrick-Lewis, D, Ciliska, D, Warren, R (2012). Screening for cervical cancer. Canadian Cervical Screening Initiative	Nil
AHRQ HTA Vesco et al. 2011	Vesco KK, Whitlock EP, Eder M, Lin J, Burda BU, Senger CA, Holmes RS, Fu R, Zuber S. Screening for cervical cancer: A systematic evidence review for the US Preventive Services Task Force. Evidence Synthesis No. 86. AHRQ Publication No. 11-05156-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; May 2011 Whitlock EP, Vesco KK, Eder M, Lin JS, Senger CA, Burda BU (2011). Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: A systematic review for the US Preventive Services Task Force. *Ann Intern Med* 155, 687–215	NTCC NETHCON

Incidence and mortality from cervical cancer (direct evidence)

No studies have assessed the impact of LBC with manual reading on incidence or mortality rates of invasive cervical cancer compared to conventional cytology. LBC is a test which will be used to identify patients at the same stage of disease as the current test used in the cervical screening program (conventional Pap tests). The present review therefore relies on evidence about the relative accuracy of manual or automated LBC to detect precancerous cervical lesions to draw conclusions about its relative effectiveness. This 'linked evidence' approach is justified by existing evidence that early detection and treatment of precancerous cervical lesions leads to a reduction in

the incidence and mortality of cervical cancer (AIHW 2012; Peto et al. 2004b, See 'Cervical cancer trends in Australia', page 10).

Similarly, the Canadian high-quality HTA (Peirson et al. 2012) did not find any studies comparing liquid-based and slide-based techniques that examined the outcomes of cervical cancer incidence or mortality (cervical cancer or all-cause) that met their inclusion criteria. The characteristics of this review are summarised in Table 61.

Detection of precancerous cervical lesions (indirect evidence)

The most recent high-quality systematic review of LBC and conventional cytology did not identify any studies eligible for inclusion based on the outcomes of cervical cancer incidence and mortality (Peirson et al. 2012). Thus the current review considers and updates evidence from the AHRQ HTA (Vesco et al. 2011) which also considered evidence on the relative accuracy of LBC and conventional cytology. A systematic literature search for LBC primary studies was conducted to identify studies that were published since the AHRQ systematic literature search, including studies from August 2011 to November 2012 to update this report. In addition, the current reviewers re-assessed studies identified in the AHRQ search but excluded from that review.

The AHRQ HTA report included two RCTs of LBC compared to conventional cytology (NTCC and NETHCON) that provided relative accuracy data and two non-randomised studies (one cohort and one split-sample) that provided absolute accuracy data for LBC and conventional cytology (CC). As RCT level evidence was identified to address this research question, studies providing a lower level of evidence were excluded.

One of the RCTs (the NTCC trial, Ronco et al. 2007a) was included in the Arbyn et al. (2008) systematic review and meta-analysis of the accuracy of LBC which was included and discussed in the previous MSAC report (2009). This RCT was also included in the current review to enable consideration of the body of level II evidence with a summary of all data presented in a consistent manner in the same report.

One additional pseudorandomised controlled trial (the RHINE-SAAR study, Klug et al. 2013) of LBC and conventional cytology was identified in the update search and included in the current review.

A master list of the primary studies included in the review of LBC effectiveness is provided in Table 45.

In addition to the AHRQ HTA another systematic review was identified (Chen et al. 2012). This was excluded as although it was considered of fair quality it excluded RCTs and split-sample studies, which was considered an inappropriate exclusion criterion.

A publication of additional analyses of unsatisfactory rate data from the NTCC and NETHCON trials (Castle et al. 2010) was also excluded; these data were included as a part of the primary publications.

Potentially included studies that were excluded from the systematic review are listed in Appendix H.

Table 45 Master list of trials included for evaluation of effectiveness of LBC versus CC for primary screening

Study identified in review	Study title	Comments
NTCC	Ronco G, Segnan N, Giorgi-Rossi P, Zappa M, Casadei GP, Carozzi F, Dalla Palma P, Del Mistro A, Folicaldi S, Gillio-Tos A, Nardo G, Naldoni C, Schincaglia P, Zorzi M, Confortini M, Cuzick J. New Technologies for Cervical Cancer Working Group. Human papillomavirus testing and liquid-based cytology: results at recruitment from the new technologies for cervical cancer randomized controlled trial. J Natl Cancer Inst, 2006; 98(11):765–774 Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Dalla Palma P, Naldoni C, Ghiringhello B, Giorgi-Rossi P, Minucci D, Parisio F, Pojer A, Schiboni ML, Sintoni C, Zorzi M, Segnan N, Confortini M. Accuracy of liquid-based versus conventional cytology: overall results of new technologies for cervical cancer screening: randomised controlled trial. BMJ, 2007; 335(7609):28. Epub May 21 2007.	Included in AHRQ Included in Arbyn et al. 2008 meta-analysis & MSAC 2009
NETHCON	Siebers AG, Klinkhamer PJ, Grefte JM, Massuger LF, Vedder JE, Beijers-Broos A. Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: A randomized controlled trial. JAMA 2009, 302(16):1757–1764	Included in AHRQ
RHINE-SAAR Study	Ikenberg H, Harlfinger W, Neis K, Konig J, Klug S. A Randomized Trial Comparing Conventional Cytology to Liquid-Based Cytology with Computer-Assistance: Results of the RHINE-SAAR Study. Journal of Cytopathology, 2011; 22(Suppl. 1):55–183 Klug SJ, Neis KJ, Harlfinger W, Malter A, König J, Spieth S, Brinkmann-Smetanay F, Kommoss F, Weyer V, Ikenberg H. A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. Int J Cancer. 2013; 132(12):2849-2857	Identified in updated literature search

Included studies

The most recent and comprehensive systematic review of the comparative accuracy of LBC and conventional cytology identified was the Agency for Healthcare Research and Quality (AHRQ) HTA conducted for the US Preventive Services Task Force (USPSTF) (Vesco et al. 2011). This review was considered a high-quality systematic review and provided data on the comparative accuracy and unsatisfactory rates of these technologies. The characteristics of this review are summarised in Table 46.

Table 46 Characteristics of most recent systematic reviews and HTAs of LBC

Author, year Country	Objective and methods	Inclusion/exclusion Included studies	Quality and applicability Conclusion
Peirson et al. 2012 Canada	Objective: Does LBC reduce the incidence of, or mortality from, invasive cervical cancer compared to CC? (Qu 1a) Literature review: • Medline, EMBASE, Cochrane • 1995–February 2011	Inclusion/exclusion criteria: *Study design:* Any *Intervention:* LBC *Comparator:* Conventional Pap *Outcomes:* Mortality, cervical ca incidence *Language:* English or French Nil studies	Quality: High Applicability: high Conclusion None
Vesco et al. 2011 AHRQ USA	Objective: Does LBC improve sensitivity, specificity and diagnostic yield and reduce indeterminate results and inadequate samples compared to CC? (Qu 2) Literature review: • DARE, HTA database, Cochrane Systematic Reviews & Trials Registry, PubMed, MEDLINE, PsychINFO, contacted authors • January 2000–September 2010 • Ancillary search 1 September 2010–3 August 2011	Inclusion/exclusion criteria: *Study design:* Systematic ref std, ref std only in positives for absolute accuracy, exclude case-control *Intervention:* LBC *Comparator:* Conventional Pap *Outcomes:* Absolute & relative test performance for detection CIN2+ *Exclusion:* Poor quality studies 4 primary studies LBC vs CC • 2 RCTs (relative accuracy) • 2 observational studies (absolute accuracy)	Quality: high Applicability: high Conclusion LBC and CC did not differ significantly in relative sensitivity or absolute sensitivity or specificity for detection of CIN2+ or CIN3+ at any cytologic threshold. LBC yielded a lower proportion of unsatisfactory slides than CC

Abbreviations: AHRQ = Agency for Healthcare Research and Quality; ca = cancer; CC = conventional cytology; LBC = liquid-based cytology; NR = not reported; Qu = question

To be included in the review, studies had to systematically apply a reference standard of colposcopy and/or biopsy. Studies were excluded if the reference standard was applied to screening test positives only (for studies of absolute test performance). Studies of automated screening technologies were excluded. Two RCTs (NTCC and NETHCON) were identified and included that reported relative test performance data, thus the measures of test performance were valid although the reference standard was not applied to all subjects. The current reviewers also extracted outcomes for the test threshold of HSIL+ from the original studies for these two pseudorandomised RCTs.

One additional primary study (Klug et al. 2013) was identified in the systematic review conducted to update the AHRQ report (Vesco et al. 2011). This pseudo-randomised study of LBC (cell filtration) versus conventional cytology provides relative sensitivities for LBC and automated image analysis of LBC (see page 102).

Study characteristics, quality and applicability

None of the included studies were conducted in HPV-vaccinated populations. The characteristics and quality of the included studies are summarised in Table 47.

The NTCC trial (Ronco 2006, 2007a) was not designed as a trial of LBC versus conventional cytology, but of HPV combined with LBC versus CC and was considered fair quality by the AHRQ authors (Vesco et al. 2011). Therefore the data presented for LBC vs CC are a secondary analysis. Whether or not patients with low-grade positive lesions were referred to colposcopy

differed between study centres, so the AHRQ authors only presented data from the sites with equivalent referral criteria for ASCUS+ thresholds. Data from both analyses are presented in the current report. The colposcopists were not blinded to the cytology or HPV test result. The recommended screening interval in Italy is three-yearly in accord with the IARC (2005) recommendations and the age range of 25–60 years is similar to the IARC (2005) recommended range of 25–65 years.

The NETHCON trial was a trial of LBC and CC and considered to be good quality by the AHRQ authors. It was cluster randomised by family practice and undertaken in the Netherlands as part of the Dutch organised cervical screening program. The recommended screening interval in the Netherlands is five-yearly.

In the RHINE-SAAR study (Klug et al. 2013), subjects were cluster randomised by week and therefore the trail is considered pseudorandomised. The study was conducted in Germany in 20 different gynaecologic practices (there is no organised screening program in Germany). In Germany women are entitled to a free annual cytological smear performed by gynaecologists. The authors state that the three-year participation rate in Germany is equivalent to the rates reached in countries with organised screening programs; however, the incidence of cervical cancer in Germany is high. Thus the population and prevalence of cervical lesions may differ to that in the Australian screening population. Patients with private health insurance were excluded; the authors state that less than 10% of women in Germany are privately insured. The authors commented that there is evidence that the sensitivity of conventional cytology in the German setting is low and that results may not be applicable to countries with high-quality organised cervical cancer screening. The applicability of this study to the Australian setting is therefore considered low. The threshold for referral to colposcopy in this study was LSIL, but verification rates at this level were low as in Germany recommended practice is repeat cytology for low-grade smears. Relative sensitivity was determined by logistic regression adjusting for verification.

Both NETHCON and NTCC trials provided data for a cytology threshold of ASCUS+ or LSIL+, but not for the threshold of HSIL+ most relevant to the Australian screening program. Relative accuracy data at this test threshold was therefore calculated by the current reviewers (for the NTCC trial this was calculated only for women ≥35 years, from Ronco et al. 2006). It should be noted that verification at the ASCUS threshold was moderate to low in the NETHCON study (Table 47), in contrast, in the NTCC trial verification was greater than 90% at the ASCUS/LSIL referral threshold.

In the study by Klug et al. (2013), the threshold for referral to colposcopy was LSIL, but verification rates at this level were low as in Germany recommended practice is repeat cytology for low-grade smears. Verification at a HSIL threshold however reached 90% (Ikenberg et al. 2011). Relative sensitivity was determined by logistic regression adjusting for verification.

Table 47 Characteristics and quality of studies reporting on the relative accuracy of LBC and conventional Pap test cytology with manual slide reading

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Ronco et al. 2006 NTCC trial Italy (3-yearly screening) 9 screening programs February 2002–June 2003	N=45,174 Inclusion criteria • Age 25–60 years • Routine screening population Exclusion criteria • Pregnancy • Hysterectomy • CIN within last 5 years Patient characteristics • Median age 41	Index test LBC cell filtration Comparator test Conventional cytology	Study design: Randomised controlled trial Test threshold: Cytology: ASCUS+ Reference standard: CIN2+ Reference standard: Colposcopy ± biopsy for CC arm ASCUS+ (7 centres) or LSIL+ (2 centres); for LBC arm ASCUS+ (or normal cytology & HPV+ if 35–60, but these results excluded from analysis). Of those referred for colposcopy, verification 93% for LBC, 91% for CC Outcomes Relative sensitivity	AHRQ included study, fair quality C1 P1 Q2 Screening study NHMRC level II

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Siebers et al. 2009 NETHCON trial Netherlands (5-yearly screening) 246 family practices April 2004–July 2006	N=89,784 Inclusion criteria • Age 30–60 years • Routine screening population Exclusion criteria NR Patient characteristics • Mean age NR	Index test LBC cell filtration Comparator test Conventional cytology	Study design: Cluster RCT: randomised by family practice Test threshold: Cytology: ASCUS+/LSIL+/HSIL+ Reference standard: CIN1+/CIN2+/CIN3+ Reference standard: Colposcopy ± biopsy for HSIL+ or persistent low grade; 18 months follow-up. Verification at HSIL+ 92% for LBC, 91% for CC; at ASCUS+ 37% for LBC, 36% for CC Outcomes Detection rates, PPV, relative sensitivity, yield	AHRQ included study, good quality C1 P1 Q1 Screening study NHMRC level II
Klug et al. 2013 RHINE-SAAR Germany (no formal screening program) 20 gynaecologic practices August 2007–March 2009	N= 20,627 Inclusion criteria • Age ≥20 years • Women presenting to gynaecologic practices Exclusion criteria • Conisation, cervical curettage • Radiation within 3 months • Hysterectomy • Private health insurance Patient characteristics NR	Index test LBC cell filtration, ThinPrep Imager Comparator test Conventional cytology	Study design: Pseudorandomised: Randomised by practice and week of appointment Test threshold: Cytology: LSIL+ Reference standard: CIN1+/CIN2+/CIN3+ (contingency data provided) Reference standard: Colposcopy ± biopsy. Verification at HSIL+ 90% Outcomes Detection rates, relative sensitivity, relative PPV	Systematic review search C1 P2 Q2 Screening study NHMRC level III-1 RCT Quality: Fair Pseudorandomised by week and practice, allocation concealment unclear Applicability limited Applicability limited as not routine screening population, no established screening program

Abbreviations: CC = conventional cytology; CIN = cervical intraepithelial neoplasia; LBC = liquid-based cytology; NR = not reported

Test accuracy

The relative accuracy measures from the included studies are presented in Table 48.

The AHRQ HTA authors concluded that there was no significant difference in the ability of LBC and conventional cytology to detect CIN2+ or CIN3+.

In contrast, the RHINE-SAAR study concluded that LBC in an opportunistic screening setting had an increased sensitivity for the detection of CIN2+ and CIN3+ at either LSIL or HSIL cytological test thresholds. These sensitivity measures were determined by logistic regression, adjusting for degree of verification. As discussed above, this study has limited applicability to the Australian screening population.

Table 48 Relative accuracy of LBC versus conventional cytology

Study	Test threshold	Relative detection rate (95% CI)	Relative PPV (95% CI)	Relative FPR (95% CI)
Detection of CIN2+	-	-	-	-
NTCC Ronco 2007a	HSIL+[e]	1.07 (0.71–1.26)[c]	0.92 (0.67–1.27)	1.42 (0.72–2.82)
	LSIL+	1.03 (0.74–1.43)	0.58 (0.43–0.78)	1.80 (1.48–2.19)
NTCC Ronco 2007a	ASCUS+	1.17 (0.87–1.56)	0.58 (0.44–0.77)	1.97 (1.75–2.21)
	ASCUS+[d]	1.11 (0.81–1.52)	0.65 (0.49–0.88)	NR
NETHCON[a] Siebers et al. 2009	HSIL+	1.03 (0.85–1.24)	1.08 (1.00– 1.16)	0.64 (0.42–0.99)
	LSIL+	1.02 (0.86–1.21)	1.04 (0.93–1.15)	0.90 (0.82–0.99)
	ASCUS+	1.01 (0.86–1.18)	1.13 (0.98–1.28)	0.90 (0.82–0.99)
Klug et al. 2013	HSIL+	3.29 (1.76–6.12)[b]	1.14 (0.79–1.65)	NR
	LSIL+	2.74 (1.66–4.53)[b]	1.17 (0.81–1.67)	NR
Detection of CIN3+	-	-	-	-
NTCC	HSIL+[e]	1.06 (0.68–1.95)	0.95 (0.52–1.38)	1.41 (0.81–2.43)
	LSIL+	0.72 (0.46–1.13)	0.40 (0.26–0.62)	1.72 (1.42–2.07)
	ASCUS+	0.84 (0.56–1.25)	0.42 (0.29–0.62)	1.93 (1.72–2.21)
NETHCON[a]	HSIL+	1.11 (0.90−1.37)	1.16 (1.02–1.32)	0.73 (0.55–0.97)
	LSIL+	1.14 (0.93−1.39)	1.17 (1.01–1.36)	0.85 (0.71–1.02)
	ASCUS+	1.09 (0.90−1.31)	1.17 (0.99–1.36)	0.89 (0.82–0.98)
Klug (2013)	HSIL+	3.94 (2.01–7.72)[b]	1.21 (0.72–2.03)	NR
	LSIL+	2.93 (1.58–5.43)[b]	1.22 (0.74–2.03)	NR

Abbreviations: CIN = cervical intraepithelial neoplasia; FPR = false positive rate; PPV = positive predictive value

[a]DRRs; [b]adjusted for verification; [c] from Arbyn et al. 2008; [d] when restricted to centres with same referral criteria, from AHRQ; [e] women ≥35 years (Ronco et al. 2006)

The MSAC (2009) report concluded that LBC provides no statistically significant difference in sensitivity (at HSIL, LSIL or pLSIL thresholds). This was based upon a meta-analysis by Arbyn et al. from 2008 which included nine primary studies (eight non-randomised studies plus the NTCC trial) that demonstrated sensitivity ratios for the detection of CIN2+ for LBC: conventional at a HSIL threshold of 1.05, 95% CI 0.95–1.16; LSIL+ 1.03, 0.96–1.11; pLSIL+ 1.03, 0.97–1.09. The most applicable data included in this review are in accord with that conclusion.

The relative false positive rate data presented from the included RCTs were inconsistent. Data from the NETHCON trial demonstrated a borderline decrease in the false positive rates with LBC, indicating a slightly increased specificity. Verification rates in this trial were low at cytological

test thresholds other than HSIL. In contrast, the NTCC trials demonstrated increased false positive rates (indicating decreased specificity) for LBC for at ASCUS or LSIL test thresholds. At the HSIL test threshold (in women 35–60 years) there was no significant difference in specificity. These data were for women aged 35 years and over in NTCC and women 30 and over in NETHCON. The NTCC data were included in the meta-analysis of relative specifities by Arbyn et al. (2008) presented in the MSAC (2009) report.

The MSAC (2009) report concluded that LBC did not differ significantly in specificity (at HSIL or LSIL thresholds) for the detection of CIN2+ compared to conventional cytology and that LBC reduces the specificity for the detection of CIN2+ at a test threshold of pLSIL. This was based upon a meta-analysis by Arbyn from 2008 including eight non-randomised studies with reference standard validation by colposcopy for all subjects. This meta-analysis demonstrated specificity ratios for the detection of CIN2+ for LBC: conventional at a HSIL threshold of 0.99, 95% CI 0.98–1.0; LSIL+ 0.97 95% CI 0.94–1.01; pLSIL (ASCUS+) of 0.91, 95% CI 0.84–0.98. The trial data presented in this update review are for women over 30–35 years and the body of evidence is considered satisfactory. The NTCC findings for relative false positive rates are in accord with these conclusions. The NETHCON trial data indicate a borderline decrease in false positive rates (relating to the inverse of specificity), but with a high potential for bias at the lower test thresholds due to low verification.

There are no significant new findings in the two included applicable RCTs that are considered to contradict MSAC's 2009 conclusions.

Unsatisfactory rates and test yield

The AHRQ review (Vesco et al. 2011) considered the unsatisfactory rates from the NETHCON and NTCC trials (N=134,162) and concluded that "LBC yielded a lower proportion of unsatisfactory slides than conventional cytology."

The test yields and unsatisfactory rates from the studies included in the current review (including NTCC and NETHCON) are presented in Table 49.

The different studies had a wide range in the unsatisfactory rates for conventional cytology. In the NTCC and NETHCON trials, the unsatisfactory rates for LBC were significantly lower than that for conventional cytology. The NTCC trial data were included in the Krahn et al. (2008) meta-analysis which was discussed in the MSAC (2009) report.

In the RHINE-SAAR study (Klug et al. 2013) the unsatisfactory rate for LBC is significantly higher than that for conventional cytology; however, the unsatisfactory rate for conventional cytology is much lower than in the Australian setting or in the other studies. This study is considered to have a low applicability to the Australian setting as it is not conducted in a country with an organised screening program.

An excluded study providing an analysis of the unsatisfactory rates from the NTCC and NETHCON trials investigating the relationship between subject age and unsatisfactory rates found no consistent relationship between age and unsatisfactory rates across the trials (Castle et al. 2010). A relationship between these factors existed within the trials, dependent upon on the particular criteria used to define unsatisfactory results.

In the NTCC trial, LBC classified significantly more slides as abnormal for all classifications (relative frequency LBC: CC ASCUS 1.57, 95% CI 1.41–1.75; LSIL 1.57, 95% CI 1.41–1.75; HSIL 1.57, 95% CI 1.13–2.18). In the NETHCON trial the classification of slides did not significantly differ between LBC and conventional cytology, although the ASCUS, AGUS and LSIL cytological categories were not differentiated (Table 49).

Based on a large meta-analysis (Krahn et al. 2008), the MSAC (2009) report concluded that LBC reduces that rate of unsatisfactory smears in comparison with conventional cytology. Findings from the two applicable studies included in the current review were consistent with these conclusions. MSAC (2009) also concluded that LBC classifies significantly more slides as positive for low-grade lesions, based upon a large meta-analysis. One of two applicable studies included in this review also found LBC had a higher rate of classification of slides as having abnormalities, across all cytological categories.

Table 49 Slide classifications of LBC versus conventional cytology slides

	LBC (%)						CC (%)					
Study	N	Unsat	pLSIL	LSIL	HSIL	AIS/SCC	N	Unsat	pLSIL	LSIL	HSIL	AIS/SCC
NETHCON	49,222	0.37**	2.07[a]	-	0.61	-	40,562	1.09	2.22[a]	-	0.63	-
NTCC	22,708	2.57**	3.59[b]	2.32	0.41	-	22,466	4.11	2.29[b]	1.26	0.26	-
Klug 2013	11,331	0.31**	2.0[c]	1.3**	1.0**	0.03	9,296	0.04	1.9[c]	0.6	0.4	0.02

Abbreviations: AIS = adenocarcinoma in situ; CC = conventional cytology; HSIL = high-grade squamous intraepithelial lesion; LBC = liquid based cytology; LSIL = low-grade squamous intraepithelial lesion; SCC = squamous cell carcinoma; Unsat = unsatisfactory

[a]ASCUS/AGUS and LSIL; [b]ASCUS/AGUS; [c]ASCUS

** P<0.001 (Chi-square)

Different LBC platforms

All included studies in this update of the MSAC (2009) review were of cell filtration LBC. No studies of cell enrichment compared with conventional cytology were of an equivalent level of evidence and therefore were not included in this update. The Arbyn et al. (2008) meta-analysis of accuracy data included in MSAC (2009) included nine studies; six were of cell filtration LBC. The Krahn et al. (2008) meta-analysis of unsatisfactory rates included in the MSAC (2009) review included studies of both cell filtration and cell enrichment. A direct comparison of the two different LBC platforms was beyond the scope of this review. Further discussion of the type of evidence available for the two different LBC platforms compared to conventional cytology is included in Appendix I.

LBC with automated image analysis for primary screening

Secondary research question 2.1

What is the comparative safety, effectiveness and cost-effectiveness of using automated image analysis under question 2 parameters, compared with the protocol proposed in the primary question?

Summary 6: Summary of effectiveness—automated image analysis

No studies have assessed the impact of automated image analysis of LBC slides on incidence or mortality rates of invasive cervical cancer compared to manual reading. One excluded study of an obsolete system for automated reading of conventional cytology found no significant difference between slide reading techniques in terms of cancer incidence or mortality.

Test accuracy

Three diagnostic accuracy studies of fair to low quality and limited applicability provided data on the comparison of automated image analysis of LBC slides to manual LBC. These studies update the MSAC (2009) review which considered evidence from four studies (three of the ThinPrep Imager system and one of the outdated AutoPap system) providing level III-2 evidence of fair quality and limited or unclear applicability.

In the UK MAVARIC study (Kitchener et al. 2011a) automated image analysis by two systems was less sensitive and more specific than manual reading of LBC for the detection of CIN2+ and CIN3+ by a histological reference standard. When individual systems were considered, the sensitivity of automated image analysis of either system was significantly lower than manual LBC for the detection of CIN2+.

A low-quality pseudorandomised trial of automated image analysis of cell filtration slides with the ThinPrep Imager system (Palmer et al. 2012) found no difference in the accuracy for classification of slides as HSIL according to an uncertain quality cytological reference standard. A partly retrospective, low-quality study of automated image analysis of cell enrichment LBC with the FocalPoint system (Wilbur et al. 2009) found an increase in the accuracy for slides to be classified as HSIL according to adjudicated cytology.

The overall body of evidence is considered poor (consisting of level III evidence which is inconsistent) and therefore conclusions on the relative accuracy of manual LBC and automated image analysis of LBC cannot be made.

MSAC (2009) concluded that automation-assisted image analysis with the ThinPrep Imager system detects as many CIN2+ lesions as conventional cytology, and may detect more. This conclusion was based upon two fair-quality level III-2 Australian studies (Davey et al. 2007; Roberts et al. 2007). The data presented in this review for the comparison to conventional cytology from two studies of the ThinPrep Imager system are either of lower quality or applicability than the studies considered in MSAC 2009, but nevertheless are in accord with this conclusion.

Unsatisfactory rates and test yields

In three studies conducted in countries with established screening programs, automated image analysis of cell filtration LBC slides yielded significantly lower unsatisfactory rates than manual slide reading of either LBC (MAVARIC, Kitchener et al. 2011; Palmer et al. 2012) or conventional slides (Halford et al. 2010). However, as the unsatisfactory rate for conventional slides is higher in the UK than in Australia, the applicability of the findings in comparison to manual LBC is uncertain.

Conclusions

- There is limited evidence regarding the effectiveness of automated image analysis when compared with manual reading of LBC.
- Automated image analysis with the ThinPrep Imager system detects as many CIN2+ lesions as conventional cytology, and may detect more (MSAC 2009).
- There is no evidence of an advantage, disadvantage or equivalence of the accuracy of the FocalPoint system compared to conventional cytology (MSAC 2009).
- Automated image analysis with the ThinPrep Imager system yields lower unsatisfactory rates than manual slide reading of conventional slides.

Table 50 Inclusion criteria for identification of studies relevant to an assessment of automated image analysis

Selection criteria	Inclusion criteria
Population	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Intervention	Automated cervical cytology image analysis for primary screening using LBC
Comparator(s)	LBC or conventional cytology
Outcomes	Accuracy (sensitivity and specificity, or the ratio of true positive to false positive findings, or the incremental rate of true positives) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS [adenocarcinoma in situ]) in women with a possible or definite HSIL cytology result Cervical cancer incidence Mortality
Reference standard	Colposcopy with biopsy for positives
Search period	2008–2012 (HTAs); 2008–2012 (primary studies)
Language	English

Further detail provided in Box 2, page 39.
Abbreviations: AIS = adenocarcinoma in situ; CIN = cervical intraepithelial neoplasia; HSIL = high-grade squamous intraepithelial lesion; HTA = health technology assessment; LBC = liquid based cytology; SCC = squamous cell carcinoma

The most recent HTA identified and included for the assessment of evidence for automated image analysis is listed in Table 51. The inclusion criteria used are listed in Table 50.

Table 51 Relevant systematic reviews and meta-analyses identified for effectiveness of automated image analysis for primary screening

Study identified in review	Study title	Studies included in this review
Peirson et al. 2012	Peirson, L., Fitzpatrick-Lewis, D, Ciliska, D, Warren, R. (2012). Screening for cervical cancer. Canadian Cervical Screening Initiative	Nil

Incidence and mortality from cervical cancer (direct evidence)

The Canadian HTA (Peirson et al. 2012) included evidence for the outcomes of cervical cancer incidence and mortality (cervical cancer or all-cause) for automated compared with manual slide reading of conventional cytology. The characteristics of this HTA in relation to this question are presented in Table 52.

This high-quality review identified one large Finnish RCT (N=503, 391) that reported on the outcomes of cervical cancer mortality and incidence (Anttila et al. 2011). This was a study of a now obsolete system of automated reading of conventional cytology (PapNet) compared with manual reading of conventional cytology in a population-based screening program, including women attending in 2003–2005. As this study did not directly address the research question for this review and the technology used is outdated, it was not included.

Table 52 Characteristics of the most recent systematic review or HTA of automated image analysis

Author, year Country	Objective and methods	Inclusion/exclusion Included studies	Quality and applicability Conclusion
Peirson et al. 2012 Canada	Objective: Does computer-assisted screening reduce the incidence of or mortality from invasive cervical cancer compared to CC? (Qu 1c) Literature review: • Medline, EMBASE, Cochrane • 1995 to February 2011	Inclusion/exclusion criteria: *Study design:* Any *Intervention:* Automated image analysis *Comparator:* Conventional Pap *Outcomes:* Mortality, cervical ca incidence *Language:* English or French Included studies: 1 RCT of automated image analysis of CC vs CC	Quality: High Applicability: Limited Comparison to CC Conclusion No difference between the two slide reading techniques on cervical cancer incidence or mortality on 4–8 year follow-up

Abbreviations: CC = conventional cytology; CIN = cervical intraepithelial neoplasia; RCT = randomised controlled trial; Qu = question

As automated image analysis is simply a different method for interpreting the same slides, it is a test which will be used to identify patients at the same stage of disease as the manual reading of LBC slides (or conventional Pap tests). The present review therefore relies on evidence about the relative accuracy of manual or automated LBC to detect precancerous cervical lesions to draw conclusions about its relative effectiveness. For this comparison, adjudicated cytology is considered a valid (but suboptimal) referenced standard (Irwig 2004).

Detection of precancerous cervical lesions (indirect evidence)

As the most recent high-quality systematic review of automated image analysis did not identify any studies of automated image analysis of LBC slides (Peirson et al. 2012), a systematic literature search for automated image analysis primary studies was conducted to identify studies that were published since the previous MSAC (2009) review, including studies from 2008 to 2012.

Automated image analysis of LBC versus manual LBC

Three studies were identified that provided data on the comparative accuracy of automated image analysis of LBC slides in comparison to manual reading of LBC slides (Kitchener et al. 2011a; Palmer et al. 2012; Wilbur et al. 2009).

The MAVARIC trial did not meet all of the inclusion criteria for this review as accuracy data for a test threshold of pHSIL/HSIL+ were not reported. Nevertheless, this study was included for review due to the paucity of evidence identified.

A Scottish government-funded feasibility study of automated image analysis reported its findings in 2012 (Palmer et al. 2012). This pseudorandomised controlled trial did not meet all of the inclusion criteria for outcomes for this review either, as only positive predictive values were reported for the outcome of histologically confirmed CIN2+ and inadequate crude data were provided to enable the determination of relative sensitivity or specificity measures. This study is nevertheless included for completeness.

The study by Wilbur et al. (2009) was a paired diagnostic accuracy study with seeded slides using adjudicated cytology as a reference standard. One arm of the study was retrospective.

Automated image analysis of LBC in comparison to conventional cytology

Two studies were identified and included that provided a comparison of automated image analysis of LBC slides to conventional cytology. One was a pseudorandomised controlled trial (the RHINE-SAAR study, Klug et al. (2013)) that provided evidence of the comparative accuracy of automated image analysis of LBC compared with manual reading of cell filtration LBC and also conventional cytology. This study was also included as providing evidence for LBC compared to conventional cytology in the previous section. A retrospective Australian study by Halford et al. (2010) reported on the performance of automated image analysis in comparison to conventional cytology.

One excluded study reported on yield and unsatisfactory rates of FocalPoint computer-assisted screening of cell enrichment liquid-based cervical cytology slides in an Australian setting (Bowditch 2011). This study did not report accuracy data, and hence did not meet the inclusion criteria for this review. Further details of this study are provided in Appendix I. Other studies excluded from the review are listed in Appendix H.

Table 53 Master list of studies included for evaluation of effectiveness of automated image analysis for primary screening

Study identified in review	Study title	Comments
RHINE-SAAR Study	Ikenberg H, Harlfinger W, Neis K, Konig J, Klug S. A Randomized Trial Comparing Conventional Cytology to Liquid-Based Cytology with Computer-Assistance: Results of the RHINE-SAAR Study. Journal of Cytopathology, 2011; 22(Suppl. 1):55–183 Klug SJ, Neis KJ, Harlfinger W, Malter A, König J, Spieth S, Brinkmann-Smetanay F, Kommoss F, Weyer V, Ikenberg H. A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. Int J Cancer. 2013; 132(12):2849-2857	Identified in updated literature search
MAVARIC	Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R, Gray A, Sadique Z, Moss S (2011). MAVARIC – a comparison of automation-assisted and manual cervical screening: A randomised controlled trial. Health Technol Assess 15, i–176 Kitchener HC, Blanks R, Dunn G, Gunn L, Desai M, Albrow R, Mather J, Rana DN, Cubie H, Moore C, Legood R, Gray A, Moss S (2011). Automation-assisted versus manual reading of cervical cytology (MAVARIC): A randomised controlled trial. Lancet Oncol 12, 56–64	Identified in updated literature search
Palmer	Palmer TJ, Nicoll SM, Mckean ME, Park AJ, Bishop D, Baker L, Imrie JEA (2012). Prospective parallel randomized trial of the MultiCyte(trademark) ThinPrep (registered trademark) imaging system: The Scottish experience. Cytopathology Scottish Cervical Cytology Review Group Feasibility Sub Group. Cervical Cytology ThinPrep Imager® (TIS) Feasibility Study – Report from the Feasibility Sub Group to Cervical Cytology Review Group. 2009. 26-3-2013	Identified in updated literature search
Wilbur	Wilbur DC, Black-Schaffer WS, Luff RD, Abraham KP, Kemper C, Molina JT, & Tench WD (2009). The Becton Dickinson FocalPoint GS imaging system: Clinical trials demonstrate significantly improved sensitivity for the detection of important cervical lesions. Am J Clin Pathol 132, 767–775	Identified in updated literature search
Halford	Halford JA, Batty T, Boost T, Duhig J, Hall J, Lee C, Walker K (2010). Comparison of the sensitivity of conventional cytology and the ThinPrep Imaging System for 1,083 biopsy confirmed high-grade squamous lesions. Diagn Cytopathol 38, 318–326	Identified in updated literature search

Included studies

In total three studies providing a comparison of automated image analysis to manual reading of LBC slides (two using cell filtration and one using cell enrichment) were included and two studies with a comparison to conventional cytology. A master list of the primary studies included in the review of automated image analysis effectiveness is provided in Table 53.

Study characteristics, quality and applicability

None of the included studies were conducted in HPV-vaccinated populations. The characteristics and quality of the included studies are summarised in Table 54.

Of the three studies providing a comparison of automated image analysis and manual LBC, the MAVARIC trial used colposcopy and biopsy as a reference standard; Wilbur et al. (2009) used adjudicated cytology as the reference standard and Palmer et al. (2012) used "final cytology" and colposcopy with biopsy as the reference standard. MAVARIC provided relative accuracy outcomes, whereas Palmer provided positive predictive values (PPVs) only for the reference standard of histology (insufficient crude data were reported to enable calculation of other outcomes), and Wilbur provided accuracy differences for the adjudicated cytology outcome.

The MAVARIC study (Kitchener et al. 2011a) provided paired data on the accuracy of LBC (cell enrichment and cell filtration) and automated image analysis using either the ThinPrep Imager or FocalPoint systems. The study is described as an RCT, with randomisation to a manual LBC-only arm and a paired arm where slides were read with manual and automation-assisted techniques. Randomisation was initially 1:1, then altered to 1:3 when accrual was below target, giving an overall 1:2. The study was powered as a non-inferiority trial. General practitioner or community clinics were also cluster randomised to either of the two automated image analysis systems – FocalPoint or ThinPrep Imager. The outcomes for the main comparison of automated image analysis to manual LBC slide reading techniques is of data from within the paired arm only, hence the data for this outcome are considered to provide the level of evidence of a paired diagnostic accuracy study, rather than an RCT. The randomised manual-only arm was designed primarily for the purpose of blinding the cytologists to whether or not slides were part of the paired reading arm. Statistical analyses were conducted to confirm that results for manual LBC tests from the paired arm of the trial were comparable to manual LBC slide reading conducted without a paired reading. While comparisons of the two automated image analysis systems were conducted, the study was not powered for these secondary analyses. The test threshold for positive cytology was the referral to colposcopy: HSIL+ or ASCUS/LSIL+ with a positive HPV triage test result. Contingency data were not provided that enable calculation of the outcomes for a pHSIL/HSIL test threshold; therefore, this study did not meet all selection criteria for this review. Outcomes were appropriately reported as relative accuracies for histological findings of CIN2+.

For a comparison of alternative methods of slide reading, the accuracy of the cytological sample for detecting precancerous cervical lesions remains unchanged. Thus comparing the ability of the reader to correctly classify the slides is a suitable surrogate to compare the performance of the technologies and adjudicated cytology is considered a valid (although suboptimal) reference standard (Irwig et al. 2004). Two studies were included that used cytology as a reference standard. Both of these studies reported the sensitivity and specificity of cytological diagnoses for detecting reference standard cytology at several reference thresholds. The current review, however, is concerned with the accuracy of different slide reading methods at detecting high-grade disease at different test thresholds, and thus accuracy for other reference thresholds has not been considered. Neither of these studies provided contingency data for the reported accuracy outcomes.

The study by Palmer et al. (2012) was designed as a feasibility study and is considered a low-quality study providing level III-1 evidence for a screening intervention. The study is conducted in a screening setting but uses the Hologic ThinPrep Remote Imaging System (MultiCyte) and therefore the applicability is considered limited. Only positive predictive values are reported for the detection of CIN2+, without contingency data. Absolute accuracy measures of primary screening were reported using 'final cytology' as the reference standard. It is not clearly described how the final cytology was determined, but it is likely that this differs to independent adjudicated cytology and that the reference standard incorporates part of the test result (and therefore that there is incorporation bias). This study is therefore considered of low quality and is only included for completeness.

Wilbur et al. (2009) reported a paired diagnostic accuracy study with seeded samples read by the BD FocalPoint GS Imaging System in comparison to cell enrichment LBC slides, including quality control (QC) rescreening in both arms. This was considered to provide level III-2 evidence for a screening technology and to be of fair quality. The non-seeded manual LBC slide results were retrospective; the automated image analysis reading of all slides and the manual reading of seeded slides was prospective. The study utilised a reference standard of adjudicated cytology. Due to the suboptimal reference standard and the retrospective component, this study was considered to have a higher risk of bias than the MAVARIC study. In previous studies, the FocalPoint GS Imaging

System has been used as classifying a percentage of slides as "not for review", however, this facility has not been used in this study. In this study the system is used to guide the screener to fields of view for all slides and also guided QC rescreening. The reference diagnoses were determined by a separate cytology adjudication centre for slides classified as abnormal, unsatisfactory or differing in classification between the study arms. For slides classified as normal in both study arms, this was taken as the reference diagnosis. The applicability of the study is limited as it contains seeded slides.

Due to the suboptimal reference standard, the study design and quality, and the lack of reported contingency data, the Palmer (2012) and Wilbur (2009) studies were considered to have a higher potential for bias for relative accuracy measures than the MAVARIC study.

The RHINE-SAAR study (Klug et al. 2013) provided evidence for a comparison of automated image analysis of cell filtration LBC slides to conventional cytology as well as for manual reading of LBC slides. The characteristics and quality of this study are discussed in detail above (see 'Manual LBC for primary screening page 94, Table 47). This study was considered a fair-quality pseudorandomised trial but is of limited applicability to the Australian setting as there is no established screening program and conventional cytology has a low sensitivity in the German setting.

An Australian study by Halford et al. (2010) was a retrospective, paired diagnostic accuracy of ThinPrep Imaging in comparison to conventional cytology. This study reported on the detection of histologically confirmed high-grade lesions in samples that had an adjunct ThinPrep specimen collected. This represented approximately 20% of the total Pap test workload of the service and therefore these samples may not be fully representative of the population as a whole. This study provided data on the number of true positive cases detected by each screening method. The proportion of positive cases that had histological verification was not reported. Data on the number of false positive cases were not reported. The study was considered a limited applicability and low-quality diagnostic accuracy study.

Table 54 Characteristics and quality of studies reporting on the relative accuracy of automated image analysis and manual slide reading

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Palmer et al. 2012 Scotland (3-yearly screening) 6 laboratories Recruitment period NR	N=169,917 (79,366 automated; 90,551 manual) Inclusion criteria • Age range NR • Screening population (20–60 years) Exclusion criteria • Nil Patient characteristics • Mean age NR	Index test Hologic ThinPrep Remote Imaging System (MultiCyte TM) Comparator test Manual LBC (ThinPrep)	Study design Parallel-group RCT, pseudo-randomised by accession number Test threshold: Cytology: HSIL Reference standard: CIN2+/high grade cytology Reference standard: Colposcopy ± biopsy for PPV; final cytology (method for determination not clear) for absolute accuracy measures Outcomes PPV for CIN2+ histology (no contingency data), yield, sensitivity, specificity for final cytology, screening rates	Systematic review search Screening study NHMRC level III-1 C1 P1 Q3 Quality: Low Inclusion criteria unclear, contingency data not provided, randomisation poor, blinding and reference standard determination unclear Applicability: Limited Screening population, use of remote automated imaging
Kitchener et al. 2011a MAVARIC trial UK (25–49 yr 3-yearly screening, 50–64 yr, 5-yearly screening) 174 GP and community clinics March 2006–February 2009	N=48,578 paired comparison arm (73,266 total) Inclusion criteria • 25–64 yr • 3,619 slides women outside screening age • Screening population • Some samples from non-randomised general practices and colposcopy clinics Exclusion criteria (N=429) • Most vault cytology or processing errors • Early repeat testing • No colposcopy result available • Inadequate samples • Urgently required samples Patient characteristics • Mean age 39 years	Index test Automated image analysis of LBC (TPI or FocalPoint) Comparator test Manual LBC (cell filtration or cell enrichment)	Study design: RCT non-inferiority trial, GP or community clinics cluster randomised by strata to automated image analysis technology, individually randomised (1:2) to manual or paired reading Test threshold: Cytology: ASCUS/LSIL+ (with HPV+ for low grade +) Reference standard: CIN2+ Reference standard: Colposcopy ± biopsy for ASCUS/LSIL with HPV triage, HSIL+ Outcomes Relative sensitivity, specificity	Systematic review search C1 P2 Q2 Screening study NHMRC level III-2 Quality: Fair Prospective accuracy study with optimal reference standard for most patients with positive result, contingency data reported Applicability: Limited Screening population, test threshold pLSIL with HPV triage

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Wilbur et al. 2009 USA 4 laboratories Recruitment from files, period NR Screening interval NR	N=12,313 Inclusion criteria • 18–75 yr • Consecutive slides • Mixed screening/diagnostic population • 361 seeded slides Exclusion criteria • 419/12,732 slides (3.3%) • Broken or cracked slides • Lack of clinical information for diagnosis • Part of multiple-slide case • Not successfully processed • Markings that could not be removed Patient characteristics • Mean age NR	Index test BD FocalPoint GS Imaging System, with guided manual screening of all slides (no NFR slides) plus directed QC rescreening Comparator test LBC cell enrichment (Sure Path) manual reading plus QC rescreening	Study design Retrospective paired diagnostic accuracy study with seeded slides Test threshold: Cytology: ASCUS+, LSIL+, HSIL+ Reference standard: Adjudicated reference diagnoses (concordant negative = negative, all others plus "some" negative slides reviewed by separate cytology adjudication centre) Outcomes Difference in sensitivity & specificity, absolute sensitivity & specificity for adjudicated cytology	Systematic review search Screening study level III-2 C1P2Q2 Quality: Fair Manual arm retrospective, not consecutive, adjudicated cytology as reference standard (valid but suboptimal) Applicability: Limited Mixed screening & diagnostic population, seeded slides
Halford et al. (2010) Australia 1 cytology service Recruitment period NR	N=87,284 Inclusion criteria • Primarily screening population • Split pair samples (approx. 20% of total) • Age range NR Exclusion criteria • NR Patient characteristics Mean age NR	Index test ThinPrep Imaging system Comparator test Conventional cytology	Study design Split-sample (paired) diagnostic accuracy study Test threshold Cytology: pHSIL, pLSIL Reference standard: high-grade lesions Outcomes Additional TPs, yield, unsatisfactory rates	Reference list checking Screening study level III-2 C1P1Q3 Quality: Low Retrospective, proportion patients with reference standard verification unclear, selection criteria unclear Applicability: Limited Selected Australian screening/diagnostic population, comparison to conventional cytology

Abbreviations: CC = conventional cytology; CIN = cervical intraepithelial neoplasia; LBC = liquid-based cytology; NR = not reported, TP = true positive

Test accuracy

Automated image analysis versus LBC manual—histological reference standard

In the MAVARIC study (Kitchener et al. 2011a), for women of all ages automated image analysis was less sensitive and more specific than manual reading of LBC for both the detection of CIN2+ and CIN3+. The relative sensitivity of automated image analysis compared to manual reading of LBC to detect CIN2+ was 0.92 (0.89 to 0.95) and the relative specificity 1.006 (1.005 to 1.007). For the detection of CIN3+ the relative sensitivity was 0.95 (0.91 to 0.99) and the relative specificity 1.007 (1.006 to 1.008).

For women in the routine screening population, relative accuracy measures were reported for the automated image analysis technologies combined and separately (Table 55). The relative sensitivity was slightly lower for FocalPoint in comparison to manual LBC than for the ThinPrep Imager in comparison to LBC, but these differences were not significantly different. It appears that there was no significant difference in the sensitivity of the ThinPrep Imager and manual LBC for the detection of CIN3+. However, the study was not powered to detect a difference at the level of individual technologies.

In the previous MSAC (2009) review, two included studies (Davey et al. 2007; Roberts et al. 2007, both using cell filtration LBC) providing a similar level of evidence and quality to that of MAVARIC found an increase in the detection of CIN2+ by automated image analysis in comparison to conventional cytology (not statistically significant in Roberts et al. 2007). Roberts et al. (2007) however provided a three-way comparison of the ThinPrep Imager, manual reading of LBC slides and conventional cytology and found that slightly fewer CIN2+ lesions were detected with automated image analysis than with manual LBC, although this trend was not statistically significant. Of two additional studies, one was consistent and one inconsistent with this trend (Bolger et al. 2006; Biscotti et al. 2005, both using cell filtration LBC). The MSAC (2009) report commented that: "It is unclear whether any increase in detection of high-grade lesions by the ThinPrep Imager system [in comparison to conventional cytology] is attributable to LBC alone, to the automation-assisted reading system, or a combination of both." The findings of MAVARIC are thus consistent with an insignificant trend in two of three included studies providing a comparison to manual LBC reading in the MSAC (2009) review.

Table 55 Relative accuracy of automated image analysis LBC versus manual LBC from MAVARIC (women aged 25–64 years, routine samples)

Automated technology	Test threshold	Relative sensitivity (95% CI)	Relative specificity (95% CI)
Detection of CIN2+	-	-	-
Combination	pLSIL+[a]	0.91 (0.87–0.95)	1.004 (1.003–1.005)
FocalPoint [b]	pLSIL+[a]	0.90 (0.85–0.96)	1.006 (1.004–1.007)
TPI[b]	pLSIL+[a]	0.92 (0.87–0.98)	1.003 (1.002–1.004)
Detection of CIN3+	-	-	-
Combination	pLSIL+[a]	0.94 (0.89–0.99)	1.005 (1.004–1.006)
FocalPoint[b]	pLSIL+[a]	0.91 (0.84–0.99)	1.006 (1.005–1.008)
TPI[b]	pLSIL+[a]	0.97 (0.91–1.03)	1.004 (1.002–1.005)

[a] borderline or mild dyskaryosis considered positive only if HPV+; [b] study not powered to demonstrate significant differences for individual technologies

Abbreviations: CIN = cervical intraepithelial neoplasia; PPV = positive predictive value; TPI = ThinPrep Imager

Automated image analysis versus LBC manual—cytological reference standard

Two studies were identified which used adjudicated or final cytology as a reference standard to compare the alternative slide reading methods. This is considered a valid (although suboptimal)

reference standard for this comparison. The accuracy data reported for this comparison are for the ability of the cytologist to correctly classify a slide according to the final classification, not for the ability of the screening technology to diagnose the cervical cancer precursor lesions. This provides informative data for the comparison but should not be considered as true accuracy values for screening for cervical cancer precursors. These outcomes are considered to provide weaker evidence for the relative accuracy of the screening technologies than the studies discussed above using histological verification as a reference standard.

The Scottish feasibility study (Palmer et al. 2012) reported no change in the positive predictive value for the detection of histologically confirmed CIN2+ or CIN3+by automated image analysis or manual LBC at high- or low-grade positive test thresholds (Table 56). Contingency data were not provided to enable the calculation of relative accuracy measures.

Table 56 Positive predictive values (95% CI) for the detection of cervical intraepithelial neoplasia in Scottish feasibility trial

Study	Test threshold	Manual	Automated	P
Detection of CIN2+	-	-	-	-
Palmer et al. 2012	HSIL+	78.5 (76.7–80.3)	80.7 (78.5–82.9)	0.140
Palmer et al. 2012	pLSIL+	51.8 (50.3–53.4)	52.8 (50.8–54.7)	0.478
Detection of CIN3+	-	-	-	-
Palmer et al. 2012	HSIL+	50.8 (48.6–53.0)	52.3 (49.5–55.1)	0.404
Palmer et al. 2012	pLSIL+	28.8 (27.4–30.2)	28.5 (26.8–30.3)	0.823

This Scottish study also provided data on the accuracy of classification of slides on primary screening according to "final cytology". The sensitivity and specificity of primary screening at a test threshold of HSIL for a final cytology classification of HSIL did not significantly differ between manual reading of LBC slides and automated reading (Table 57). Automated image analysis reading of cell filtration LBC had a sensitivity of 79.9% (95% CI 77.6 –82.3) and a specificity of 99.6% (95% CI 99.6–99.7) compared to manual reading of cell enrichment LBC slides with a sensitivity of 81.7% (95% CI 79.5–83.8) and a specificity of 99.6% (95% CI 99.5–99.6).

The study by Wilbur et al. (2009) reported on the accuracy of classification of slides according to the reference standard of adjudicated cytology (Table 57). In this study, automated image analysis reading of cell enrichment LBC had a sensitivity of 85.3% and a specificity of 95.1% compared to manual reading of cell enrichment LBC slides with a sensitivity of 65.7% and a specificity of 97.7% (95% CI not reported). This equates to a 19.6% (12.7%–26.8%; P<0.001) higher sensitivity and 2.6% (12.7%–26.8%; P<0.001) lower specificity of the automated image analysis. Given the high potential for bias in this study (retrospective component, seeded slides and suboptimal reference standard), and the lack of both contingency data and a comparison to conventional cytology, it is difficult to interpret these results.

Table 57 Relative accuracy of automated versus manual LBC for classification of HSIL (cytology reference standard)

Study	Test threshold	Sensitivity difference[a] %	Sensitivity difference[a] (95% CI)	Specificity difference[a] %	Specificity difference[a] (95% CI)
Palmer et al. 2012	HSIL+	–1.8[b]	-	0.0[b]	-
Wilbur et al. 2009	HSIL+	19.6	(12.7–26.8)	–2.6	(–3.4 to –1.9)
Wilbur et al. 2009	LSIL+	–0.5[b]	-	NR	-

Abbreviations: CI = confidence interval

* P<0.0001; [a] automated – manual; [b] calculated by reviewers as difference in reported % accuracy values

Automated image analysis versus conventional cytology

The RHINE-SAAR study (Klug et al. 2013) reported relative accuracy measures determined by logistic regression adjusting for verification (Table 58). In this study automated image analysis had a significantly greater sensitivity for the detection of CIN2+ and CIN3+ than conventional cytology. Similarly, manual LBC has a significantly greater sensitivity than conventional cytology, a finding which is discordant with the results of other RCTs (see 'Manual LBC for primary screening page 98). However, the applicability of this study is greatly limited. The authors have noted that conventional cytology has a low sensitivity in the German setting.

Table 58 Estimates of the relative accuracy of automated image analysis LBC versus conventional cytology (ITT analysis)

Study	Test threshold	Relative sensitivity Auto:CC (95% CI)	Relative sensitivity LBC:CC (95% CI)	Relative PPV Auto:CC (95% CI)
Detection of CIN2+	-	-	-	-
Klug et al. (2013)	HSIL+	3.53 (1.92–6.48)[a]	3.29 (1.76–6.12)[a]	1.19 (0.84–1.69)
Klug et al. (2013)	LSIL+	3.17 (1.94–5.19)[a]	2.74 (1.66–4.53)[a]	1.07 (0.75–1.53)
Detection of CIN3+	-	-	-	-
Klug et al. (2013)	HSIL+	4.02 (2.07–7.81)[a]	3.94 (2.01–7.72)[a]	1.21 (0.73–2.00)
Klug et al. (2013)	LSIL+	3.45 (1.88–6.34)[a]	2.93 (1.58–5.43)[a]	1.09 (0.66–1.80)

Abbreviations: ITT = intention-to-treat, PPV = positive predictive value

[a] Adjusted for verification

The retrospective Australian study by Halford et al. (2010) reported that within 87,284 Pap tests collected with an adjunct sample, the ThinPrep Imager detected 18 more true positive cases for histological high-grade disease than conventional cytology at a threshold of pHSIL (0.2 per 1,000 women screened); and 48 additional true positives at a threshold of pLSIL (0.5 per 1,000 women screened). This study was considered a low-quality diagnostic accuracy study and the potential for bias is considered high.

MSAC (2009) concluded that automation-assisted image analysis with the ThinPrep Imager system detects as many CIN2+ lesions as conventional cytology, and may detect more. This conclusion was based upon two fair-quality Australian studies of the ThinPrep Imager system (Davey et al. 2007; Roberts et al. 2007). One study demonstrated a statistically significant increase in detection of CIN2+ (Davey et al. 2007); the other demonstrated a non-statistically significant increase in detection of high-grade lesions (Roberts et al. 2007). The data presented in this review for the

comparison to conventional cytology are either of low quality or applicability, but nevertheless are in accord with this conclusion.

Unsatisfactory rates and test yield

The MSAC 2009 report concluded that LBC and automated image analysis of cell filtration slides (ie with the ThinPrep Imager system) reduced the rate of unsatisfactory smears in comparison to conventional cytology.

The rates of unsatisfactory slides and test yields from the studies included in this review are provided in Table 59. The rate of unsatisfactory slides for automated image analysis of cell filtration was significantly less than for manual LBC or conventional cytology in all studies except the low applicability German RHINE-SAAR study. However, as the unsatisfactory rate for conventional slides is higher in the UK than in Australia, the applicability of the findings in comparison to manual LBC, based on two UK studies (MAVARIC, Palmer et al. 2012) is uncertain. In the one study of automated image analysis of cell enrichment slides (Wilbur et al. 2009), unsatisfactory rates were not significantly different between automated image analysis and manual reading.

MSAC (2009) also concluded that automated image analysis of cell filtration slides (with the ThinPrep Imager system) increased the number of slides classified as having low-grade lesions on cytology in comparison to conventional cytology. The only included study reporting on this comparison in a country with an established screening program (Australia) is in accord with this conclusion (Halford et al. 2010). In a study conducted in Germany, where there is no established screening program, the ThinPrep Imager system classified significantly fewer slides as LSIL, and significantly more slides as HSIL in comparison with conventional cytology.

In the current review, the included studies reported test yields separated by different categories, making comparisons difficult. Two studies reported significantly more lesions were classified as low-grade positive by automated image analysis of cell filtration LBC (Palmer et al. 2012; Klug et al. 2013). The other two studies reported that automated image analysis classified significantly less slides as positive for low-grade lesions (MAVARIC; Wilbur et al. 2009 for ASCUS slides only). Of the latter two studies one was a study of both image analysis systems combined, the other was a study of FocalPoint system with automated image analysis of cell enrichment slides. The inconsistencies in these data make interpretation difficult.

Different automated image analysis systems

No studies on the FocalPoint system were included in the MSAC 2009 report. MSAC (2009) included three studies of the AutoPap system (an outdated version of the FocalPoint system) in comparison to conventional cytology (Confortini et al. 2003; Stevens et al. 2004) or manual LBC (Wilbur et al. 2002).

In the current update review no studies comparing automated image analysis with the FocalPoint system to conventional cytology were included. No studies of FocalPoint image analysis of cell enrichment LBC compared with conventional cytology were of an equivalent level of evidence to those of the ThinPrep Imager system.

For a comparison to manual LBC, MSAC (2009) included: three studies comparing the ThinPrep Imager system with manual reading of cell filtration slides (Roberts et al. 2007; Bolger et al. 2006; Biscotti et al. 2005) and one study of the AutoPap system compared with manual reading of cell enrichment slides (Wilbur et al. 2002). The current update review includes two level III-1 screening studies of the ThinPrep Imager system (Palmer et al. 2012, N=79,366; Klug et al. 2013, N=20,627), one fair quality level III-2 screening study of limited applicability that provided a comparison of both automated image analysis systems to manual LBC (N=45, 578; MAVARIC, Kitchener et al. 2011a) and one level III-2 study with seeded samples of the BD FocalPoint GS Imaging System (N=12,313; Halford et al. 2010).

The MAVARIC study was underpowered for a direct comparison of the two systems, and such a comparison is considered beyond the scope of this review.

Overall, the body of evidence available for conclusions regarding the FocalPoint system is smaller and of a lower level of evidence than that available for the ThinPrep Imaging system.

Table 59 Slide classifications of automated image analysis of LBC versus conventional cytology slides

-	-	Auto LBC (%)									Manual reading (%)							
-	Study	N	Not read	Unsat	pLSIL	LSIL	pHSIL	HSIL	AIS/SCC	Gland	N	Unsat	pLSIL	LSIL	pHSIL	HSIL	AIS/SCC	Gland
Auto LBC vs LBC (cell filtration)	MAVARIC	48,271	3.56	1.82**	4.2[e]**	-	-	1.2	-	-	48,271	2.8	5.5[e]	-	-	1.3	-	-
-	Palmer et al. 2012	79,366	3.55	1.9**	8.0[a]**	-[a]	-	1.5[b]	-	-	90,551	2.7	7.5[a]	-[a]	-	1.4[b]	-	-
-	Wilbur et al. 2009	12,313	NR	0.22	3.1**	7.7[c]*	-	-	-	-	12,313	0.16	4.8	6.9[c]	-	-	-	-
-	RHINE-SAAR	11,331	NR	2.86	2.1	1.8*	-	1.0	0.02	-	11,331	0.31	2.0	1.3	-	1.0	0.03	-
Auto LBC vs CC	RHINE-SAAR	11,331	NR	2.86**	2.1	1.8**	-	1.0**	0.02	-	9,296	0.04	1.9	0.6	-	0.4	0.02	-
-	Halford 2010	87,284	NR	0.87**	4.41**	2.82**	0.65	0.97	0.04	0.20[d]	87,284	3.65	3.61	2.43	0.63	0.94	0.04	0.21[d]

Abbreviations: AIS = adenocarcinoma in situ; CC = conventional cytology; Gland = glandular; LBC = liquid-based cytology; NR = not read; SCC = squamous cell carcinoma; Unsat = unsatisfactory

[a] ASCUS & LSIL; [b] HSIL+; [c] LSIL+; [d] excluding adenocarcinoma; [e] pLSIL/LSIL

* P<0.05; ** P<0.001

HPV testing to triage women with pLSIL/LSIL

Primary research question 2.2

What is the comparative safety, effectiveness and cost-effectiveness of adding an HPV test to triage women with pLSIL/LSIL, compared with the protocol proposed in the primary question (manual LBC)?

Summary 7: Summary of effectiveness—HPV testing to triage women with pLSIL/LSIL

No studies have assessed the impact of HPV triage in women with pLSIL or LSIL cytology on incidence and mortality rates from invasive cervical cancer in comparison to repeat cytology. Therefore, a modelled analysis of cervical cancer screening, diagnosis and treatment using the IARC recommendations for HPV-vaccinated and unvaccinated populations is in progress to explore the potential long-term benefits and harms of these triage options in the Australian setting.

Detection of precancerous cervical lesions

In three RCTs, there was no significant difference in the overall detection of CIN3+ by HPV triage in comparison to either repeat cytology or immediate colposcopy.

In the ALTS trial, there were fewer CIN2 lesions detected in women referred with pLSIL in the repeat cytology arm compared to both HPV triage and immediate colposcopy. Similarly, in the Bjerre et al. trial (2008), there were fewer CIN2 lesions detected in women younger than 35 in the repeat cytology arm compared to the HPV triage arm (but not in women over 35). These data suggest that a statistically significant proportion of prevalent CIN2 lesions in trial participants, particularly younger women, regressed when a strategy of repeat cytology was followed, potentially avoiding unnecessary treatment and associated harms.

The colposcopy rate was higher in the HPV triage arms than the repeat cytology arms of trials and in women less than 35 years compared to older women.

Test accuracy

MSAC (2009) concluded that "if index cytology shows pLSIL, the HPV triage test is more sensitive than a single repeat cytology test for the detection of CIN2+ lesions and has similar specificity".

The pooled accuracy data from Arbyn (2013) are in accord with this conclusion and the same results were seen for the detection of CIN3+ lesions.

MSAC (2009) also concluded that "if index cytology shows LSIL, the HPV triage test is not more sensitive than a single repeat cytology test for the detection of CIN2+ lesions and has lower specificity".

The pooled accuracy data from Arbyn et al. (2013) are in accord with this conclusion regarding specificity and the same finding was made for the detection of CIN3+.

The Arbyn et al. (2013) meta-analysis found that the HPV triage test for LSIL is more sensitive than a single repeat cytology test for the detection of CIN2+ lesions, but not CIN3+ lesions. Whether or

not any additional CIN2+ lesions detected are destined to progress or regress cannot be determined from these data.

Age subgroups

MSAC (2009) concluded that "restricting the HPV triage test to older age groups is associated with a higher specificity and lower colposcopy referral rate and a corresponding smaller gain in sensitivity compared with its use in all age groups."

The pooled accuracy data from Arbyn et al. (2013) found that the sensitivity of HC2 triage did not vary significantly by age but the specificity was always increased with age (three studies). Data from one included RCT demonstrated a higher colposcopy rate in women aged <35 compared to women aged ≥35 years following HPV triage.

Conclusions

- The HPV triage test is more sensitive than a single repeat cytology test for the detection of CIN2+ and CIN3+ lesions in women with pLSIL, and has similar specificity.
- The HPV triage test is more sensitive than a single repeat cytology test for the detection of CIN2+ lesions (but not CIN3+) in women with LSIL, and has lower specificity.
- A significant proportion of additional CIN2+ lesions that would be detected by HPV triage of pLSIL and LSIL are likely to regress when a strategy of repeat cytology is used.
- The colposcopy rate following HPV triage is higher in women aged <35 years compared to women aged ≥35 years.

MSAC last considered evidence for HPV testing to triage women with pLSIL or LSIL in 2009. The review included the ALTS RCT (ALTS Group 2003a, 2003b) and a 2006 meta-analysis (Arbyn et al. 2006). The review concluded that:

- An immediate HC2 HPV triage test is a more sensitive test than a single repeat cytology test for detecting CIN2+ lesions in women with pLSIL, and has similar specificity.
- An immediate HC2 HPV triage test is no more sensitive than a single repeat cytology test for detecting CIN2+ lesions in women with dLSIL, and has lower specificity; but colposcopy referral rates may be favourable compared with a strategy of two annual repeat cytology tests.
- Restricting the HPV triage test to older age groups is associated with a higher specificity and lower colposcopy referral rate and a corresponding smaller gain in sensitivity compared with its use in all age groups.
- No published clinical studies comparing the accuracy of performing the HC2 HPV test with repeat cytology at the 12-month follow-up visit with the current strategy of performing repeat cytology alone were identified.
- No published clinical studies comparing the impact of the HPV triage test on cervical cancer incidence or mortality with repeat cytology were identified.

Included studies

The most recent systematic reviews and HTAs identified and included for the assessment of evidence for HPV triage of LSIL/pLSIL are listed in Table 60.

Table 60 Relevant systematic reviews and meta-analyses included for effectiveness of HPV triage for women with pLSIL/LSIL cytology

Study identified in review	Study title	Studies included in this review
Arbyn et al. 2013	Arbyn M, Roelens J, Simoens C, Buntinx F, Paraskevaidis E, Martin-Hirsch PPL, Prendiville WJ (2013). Human Papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions (Review). Cochrane Database of Systematic Reviews, Issue 3	39 studies for accuracy of ASC-US triage 24 studies for accuracy of LSIL triage
AHRQ HTA, Vesco et al. 2011	Vesco KK, Whitlock EP, Eder M, Lin J, Burda BU, Senger CA, Holmes RS, Fu R, Zuber S. Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 86. AHRQ Publication No. 11-05156-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; May 2011 Whitlock EP, Vesco KK, Eder M, Lin JS, Senger CA, & Burda BU (2011). Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 155, 687–215	ALTS Bjerre et al. 2008

The most recent high-quality systematic review identified which addresses the research question was Arbyn et al. 2013; a Cochrane review published after the dates of our literature search was identified though contact with clinical experts. This review was included as it was considered a more recent and comprehensive update of the Arbyn et al. 2012 meta-analysis identified in the HTA search and is considered the highest quality review available. Nevertheless, the Arbyn et al. 2013 review has accuracy as the primary outcome. Therefore, the AHRQ HTA (Vesco et al. 2011) is also included to provide information on the longitudinal outcomes of randomised trials including the detection of precancerous cervical lesions and the referral rates to colposcopy. The AHRQ HTA included two RCTs: ALTS (2003 a&b), which was also included in the 2009 MSAC review and Bjerre et al. (2008).

The AHRQ HTA conducted for the US Preventive Services Task Force (USPSTF) (Vesco et al. 2011) was considered a high-quality systematic review. The characteristics of this review are summarised in Table 61. A literature search was conducted to update the AHRQ HTA.

For studies evaluating HPV triage testing, Vesco et al. (2011) included studies that provided evidence on the absolute and relative test performance, applied a reference standard of colposcopy and/or biopsy to all those screening positive and at least a random sample of those screening negative, and were conducted in routine screening populations. The review considered evidence to September 2010 and conducted an ancillary search which identified additional studies published to August 2011.

The Arbyn et al. review (2013) included studies of women presenting with a cervical cytology result of ASCUS or LSIL, who had undergone both HC2 testing and repeat cytology, or HC2 testing alone, and were subsequently subjected to reference standard verification with colposcopy and colposcopy-directed biopsies for histologic verification. The literature was searched to January 2011. This Cochrane review was considered high quality.

Table 61 Characteristics of most recent systematic reviews and HTAs of HPV triage of pLSIL/LSIL cytology

Author, year Country	Objective and methods	Inclusion/exclusion Included studies	Quality and applicability Conclusion
Vesco et al. 2011 AHRQ USA	Objective: What are the benefits of using HPV testing as a screening test, either alone or in combination with cytology, compared to conventional cervical cytology (KQ3)? Literature review: • DARE, HTA database, Cochrane Systematic Reviews & Trials Registry, PubMed, MEDLINE, PsychINFO, contacted authors • January 2000–September 2010 • Ancillary search 1 September 2010–3 August 2011	Inclusion/exclusion criteria: *Study design:* Systematic ref std, ref std only in positives for absolute accuracy, exclude case-control *Intervention:* HPV (HC2 or PCR) alone or in combination with cytology *Comparator:* Conventional or liquid based cytology *Outcomes:* Absolute & relative test performance for detection CIN2+ *Exclusion:* Poor quality studies, non-English language Reflex HPV: 2 RCTs, 4 cohort studies	Quality: High Applicability: High Conclusion For the triage of women with ASC-US cytology to colposcopy, a single HC2 test has a higher sensitivity and similar specificity compared to single repeat cytology at a threshold of ASC-US for the detection of CIN2+. No additional benefit occurs when HC2 triage is combined with cytology, but this strategy increases false positives. HC2 does not appear useful for the triage of women with LSIL cytology because such a high proportion of women will test positive.
Arbyn et al. 2013 Cochrane	Objective: To compare the accuracy of HPV testing with HC2 to that of repeat cytology for detection of underlying CIN2+ or CIN3+ in women with ASCUS or LSIL Literature review: • Cochrane Register of Diagnostic Test Accuracy Studies, CENTRAL, MEDLINE, EMBASE, hand searching of select journals • Search 1992–January 2011	Inclusion: *Study design:* studies with concomitant testing, or HC2 alone, followed by verification with the reference standard RCTs where study participants were randomised to HPV-based triage or repeat cytology and subsequently all women were submitted to verification with the reference standard *Intervention:* HC2 *Comparator:* Repeat cytology (CC or LBC) between 3 and 12 months. *Test threshold:* Cytology ASCUS+, LSIL+ and HSIL+. *Outcomes:* Histologically confirmed CIN2+ and CIN3+. 39 studies of HC2 triage accuracy in women with ASCUS (10 with repeat cytology) 24 studies of HC2 triage accuracy in women with LSIL (6 with repeat cytology)	Quality: High Applicability: High Conclusion HPV triage with HC2 can be recommended to triage women with ASCUS because it has higher accuracy (significantly higher sensitivity and similar specificity) than repeat cytology. When triaging women with LSIL, an HC2 test yields a significantly higher sensitivity, but a significantly lower specificity, compared to repeat cytology. Therefore, practice recommendations for management of women with LSIL should be balanced, taking local circumstances into account.

Abbreviations: ca = cancer; CC = conventional cytology; LBC = liquid-based cytology; NR = not reported; KQ = key question

The literature search to update the AHRQ HTA (Vesco et al. 2011) identified one additional RCT, a cluster randomised trial of 3,319 women participating in organised cervical screening in Sweden (Dillner et al. 2011). A second trial, HPV FOCAL (Ogilvie et al. 2012), comparing HPV triage of LBC with LBC triage of HPV was also identified but this comparison is not directly applicable to the research question addressed here and therefore this study is not considered further. This trial is considered in the section on HPV primary screening with LBC triage (page 168). A list of the included trials is provided in Table 62.

Table 62 Master list of trials included for evaluation of effectiveness of HPV triage of women with pLSIL/LSIL cytology with outcomes of detection of precancerous cervical lesions

Study identified in review	Study publications (publications identified in the update search in bold)	Comments
ALTS (2003a, 2003b)	ASCUS-LSIL Triage Study (ALTS) Group. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. Am J Obstet Gynecol. 2003;188:1393–1400 ASCUS-LSIL Triage Study (ALTS) Group. Human papillomavirus testing for triage of women with cytologic evidence of low-grade squamous intraepithelial lesions: baseline data from a randomized trial. J Natl Cancer Inst. 2000;92:397–402 ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. Am J Obstet Gynecol. 2003;188:1383–1392 Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study: design, methods and characteristics of trial participants. Acta Cytol. 2000;44:726–742 Sherman M, Schiffman M. Effects of age and human papilloma viral load on colposcopy triage: data from the randomized Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS). J Natl Cancer Inst. 2002;102–107 Solomon D, Schiffman M, Tarone R; ALTS Study Group. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst. 2001;93:293–299	Included in AHRQ & MSAC 2009
Bjerre et al. 2008	Bjerre P, Silfverdal L, Dillner L, et al. A randomized trial of basing treatment on human papillomavirus and/or cytology results in low-grade cervical lesion triage. Am J Obstet Gynecol. 2008;199:24–27	Included in AHRQ
Dillner et al. 2011	**Dillner L, Kemetli L, Elfgren K, et al. Randomized health services study of human papillomavirus based management of low-grade cytological abnormalities. Int J Cancer. 2011;129:151–159**	New trial identified in updated literature search.

Study characteristics, quality and applicability

Of the three included trials, the intervention arm in one was LBC with HPV triage (ALTS Group 2003a, 2003b) and two were CC with HPV triage (Bjerre et al. 2008; Dillner et al. 2011). The control arms of the three trials were repeat cytology in ALTS (LBC) and Bjerre (CC) and colposcopy in Dillner. The control arm in the Dillner trial does not directly address the research question, reducing its applicability.

All studies used the HC2 HPV test and were conducted in developed countries with established screening programs.

Two of the studies (Bjerre et al. 2008; Dillner et al. 2011) referred all women with ASCUS cytology to colposcopy in the control arm of the trial (Bjerre after repeat cytology) which is a lower threshold than used in Australia. The ALTS trial more closely resembles the Australian algorithm with repeat cytology for low-grade abnormalities and referral to colposcopy for HSIL+ cytology. In the ALTS trial all women received colposcopy at two years.

The AHRQ HTA (Vesco et al. 2011) rated both the ALTS study and Bjerre as good quality. The key quality issue for ALTS was applicability in the American context; however, this study has higher applicability in the Australian context. The key quality issue for Bjerre was the small sample size and that results are not reported separately for ASCUS and LSIL referrals. The study by

Dillner was considered fair quality; the applicability to this review is limited as the control arm did not undergo repeat cytology.

The characteristics of the RCTs identified to address this question are described in Table 63.

Table 63 Characteristics and quality of studies reporting on the relative accuracy of triage of pLSIL/LSIL with HPV compared to CC or LBC

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality & applicability
Dillner et al. 2011 Sweden 15 outpatient clinics March 2003–January 2006	N=3,319 Inclusion criteria Routine screening population, all women included Exclusion criteria None Patient characteristics Mean age: NR	Index test Colposcopy and biopsy of all women with ASCUS or CIN1 Comparator test HPV-based triaging, referring all women with ASCUS or CIN1 for a new visit with HPV testing. HPV+ women referred for colposcopy and HPV- women for repeat cytology at 12 months	Study design: Cluster-randomised by outpatient clinic Test threshold: Cytology: ASCUS+/CIN1+ HPV HC2: 1.0 RLU Reference standard: Colposcopy ± biopsy with follow up cytology at 12 and 24 months post colposcopy.	Quality: Fair Applicability: • Not a direct comparison to repeat cytology • Applicable population and setting
ALTS USA 4 centres Recruitment period: 1997–1998 NOTE: HPV triage arm closed to LSIL referrals in first year because majority of women tested positive.	N=5,060 Inclusion criteria: Community read cytologic diagnosis of ASC-US or LSIL within six months of enrolment, age ≥18 years, able to provide informed consent, likely to participate for full duration of trial Exclusion criteria: Prior hysterectomy, history of ablative or excisional therapy to cervix, pregnant Patient characteristics: Mean age: pLSIL 29 y/ LSIL 25 y Range: NR (≥ 18 y)	Index test Repeat LBC at enrolment, 6, 12 and 18 months Comparator test HCII (13 HR HPV types) + immediate colposcopy if rpt cytology showed HSIL+ Both tests performed at the time of study entry ~ 2 months after index cytology finding	Study design: RCT, 3 arms: HPV triage, immediate colposcopy, rpt cytology with 2-year follow-up Test threshold: Cytology: HSIL (or pLSIL at 2 years). Analysis for cytology threshold = pLSIL also reported HPV HCII: HPV = 1 pg HPV DNA/mL or missing/insufficient specimen Reference standard: Colposcopy ± biopsy or endocervical curettage and 2 year follow-up. Colposcopy performed immediately (colposcopy study arm) or following a positive HPV or cytology finding (comparator test arms). Pelvic examination, cytology, masked HPV testing, cervicography every 6 months with clinical review of all results and colposcopy at two-year follow-up for all women.	AHRQ Rated GOOD - Repeat cytology threshold of HSIL for referral colposcopy versus ASC-US in recent (US) guidelines. (Note that Australian guidelines refer at a HSIL threshold.)

Bjerre et al. 2008 Sweden 2 Counties Recruitment period: Varmland county: August 2001–February 2003 Orebro county: September 2002–October 2003	N=674 Inclusion criteria: Age 23–60yr (invitation to cervical screening program) Exclusion criteria: Pregnant or treated for dysplasia in last two years. Patient characteristics Mean age: 36.7 years	Index test Repeat Pap only Comparator test Repeat Pap and HPV test	Study design: RCT of women positive for ASC-US or LSIL Test threshold: Cytology: ASC-US+ HPV HCII: HC 2 ≥ 1 pg/ml Reference standard: Women with positive repeat screening tests treated with LEEP, laser conisation or hysterectomy (procedures which also provide tissue for histology) regardless of colposcopy findings.	AHRQ Rated GOOD - Small sample size

Abbreviations: ALTS = ASCUS/LSIL Triage Study; ASCUS = atypical squamous cells of undermined significance; pLSIL = possible low-grade squamous intraepithelial lesion; LSIL = low-grade squamous intraepithelial lesion; NR = not reported; RCT = randomised controlled trial

Incidence and mortality from cervical cancer (direct evidence)

This review did not identify any studies designed to compare rates of invasive cervical cancer incidence or mortality in women who received an HPV triage test as a replacement for or addition to repeat cytology for the management of LSIL abnormalities.

Detection of precancerous cervical lesions (indirect evidence)

Three RCTs compared the cumulative incidence of precancerous cervical lesions among women with low-grade cytological abnormalities assessed by HPV triage versus repeat cytology or immediate colposcopy. Results are reported in all these trials as CIN2 rather than CIN2+ which differs from HPV-based screening trials in 'Results of assessment – effectiveness: HPV testing (scenario 3)', page 137. Results from the largest of the trials, ALTS, are reproduced from MSAC 2009 and are presented in Table 64. Discussion of these results is reproduced from MSAC 2009 in the box below.

MSAC 2009, pp. 32–33

This trial did not find a statistically significant difference in the risk of CIN3 at two years for women with pLSIL or dLSIL assessed by HPV triage, immediate colposcopy or repeat cytology alone. Cumulative diagnosis rates for CIN2 at two years were lower among women with pLSIL assigned to repeat cytology than among women assigned to HPV triage or immediate colposcopy (p=0.005). This difference was also observed among women with dLSIL, although it reached statistical significance only for the comparison between repeat cytology and immediate colposcopy. Enrolment of women with dLSIL into the HPV triage study arm closed early, and the number enrolled (N=224) may have been too few to detect a true difference in CIN2 diagnosis rates between HPV triage and repeat cytology or immediate colposcopy.

These data suggest that a statistically significant proportion of prevalent CIN2 lesions in trial participants with a cytological finding of pLSIL or dLSIL regressed over two years when a strategy of repeat cytology was followed, potentially avoiding unnecessary treatment. However, this strategy included repeat cytology every six months and referral to colposcopy for a finding of HSIL at 6, 12 or 18 months or of persistent low-grade cytological abnormalities at two years, whereas current Australian guidelines recommend colposcopy if repeat cytology at 12 or 24 months shows pLSIL, dLSIL or HSIL. Thus, these trial results may not apply to the current strategy of repeat cytology in Australia.

Referral to colposcopy within the a priori-defined periods for each strategy (enrolment period for HPV arm and enrolment plus follow-up period for repeat cytology) was significantly higher in the HPV triage arm than the repeat cytology arm for both pLSIL (55.6% vs 12.3%, p=<0.001) and LSIL women (85.3% vs 18.8%, p=<0.001).

Table 64 Two-year cumulative diagnosis rate of CIN2–3 in women with an index cytology finding of pLSIL or LSIL managed by HPV triage, immediate colposcopy or repeat cytology, ALTS trial (N= 5,060) (from MSAC 2009)[1]

Index cytology N	2-year cumulative risk of CIN2 -	-	-	-	2-year cumulative risk of CIN3 -	-	-	-
	HPV triage	Immediate colposcopy	Repeat cytology	Total (test of comparison)[2]	HPV triage	Immediate colposcopy	Repeat cytology	Total (test of comparison)[2]
pLSIL 3,488	7.3%	7.9%	4.7%	6.7% (p=0.005)	8.7%	9.3%	8.3%	8.8% (p=0.72)
LSIL 1,572	10.7%	13.4%	7.6%	10.5% (p=0.002)[3]	18.3%	15.2%	13.8%	15.0% (p=0.26)

Abbreviations: LSIL = low-grade squamous intraepithelial lesions; pLSIL = possible low-grade squamous intraepithelial lesions
1. All trial participants received repeat cytology at 6, 12 and 18 months and were referred to colposcopy if HSIL. CIN2–3 detection based on results from exit colposcopy at two years and clinical review of all study findings (ALTS 2003a, 2003b); 2. χ^2 test of comparison between study arms; 3. No statistically significant difference for comparison repeat cytology vs HPV triage, p<0.001 for comparison repeat cytology vs immediate colposcopy

Results from the small Swedish trial (Bjerre et al. 2008) included in the AHRQ report (Vesco et al. 2011) are presented in Table 65. In this trial, after one round of triage, the detection of CIN2 and CIN3+ was greater in the HPV triage arm but this only reached significance for women younger than age 35 for the detection of CIN2. Unlike in the ALTS trial in which all women received an exit colposcopy, only women with positive HPV or cytology were referred to colposcopy. However, the referral threshold was low (ASC-US) in this trial. The significant increase in CIN2 detection in younger women in this small trial suggests an excess detection of regressive lesions in the HPV triage arm in these women.

Colposcopies were more frequent in those in the HPV triage arm than the repeat cytology arm (HPV triage: 61.7% (56.4–66.8) vs repeat cytology 41.8% (36.7–47.2) RR: 1.48 (1.27–1.71, p=<0.001)) and in younger women than older women (HPV triage age <35: 70.9% (63.6–77.3) vs HPV triage age ≥35: 52.9% (45.5–60.2) RR 1.34 (1.13–1.59, p=<0.001). The relative false positive proportion for the detection of CIN3+ was 1.74 (1.38–2.20) overall. Age-specific results suggest a higher relative false positive rate for women less than 30 (<30: 1.94 (1.43–2.65) vs ≥35: 1.42 (1.01–2.01)).

Table 65 Cumulative diagnosis rate of CIN2–3 in women with an index cytology finding of pLSIL or LSIL managed by HPV triage or repeat cytology, Bjerre et al. trial (N=674)[1]

Age N	Cumulative risk of CIN2 -	-	-	Cumulative risk of CIN3+ -	-	-
	HPV triage	Repeat cytology	Total (test of comparison)[2]	HPV triage	Repeat cytology	Total (test of comparison)[2]
<35 340	16.4%	8.0%	12.1% (p=0.028)	24.2%	22.3%	23.2% (p=0.77)
≥35 334	7.6%	6.8%	7.2% (p=0.95)	18.6%	13.0%	15.9% (p=0.21)
Total 674	11.9%	7.4%	9.6% (p =0.068)	21.4%	17.8%	19.6% (p=0.29)

1. Only trial participants who were HPV+ or ASCUS+ were referred to colposcopy. 2. χ^2 test of comparison between study arms

Results from a large cluster randomised trial conducted within the Swedish-organised screening program are presented in Table 66. In this trial the control arm was immediate colposcopy rather than repeat cytology. There are no statistically significant differences between the control arm and the HPV triage arm in the detection of CIN3+. There is a marginally higher detection of CIN2 in the HPV triage arm for women referred with LSIL and overall which may be due to higher colposcopy compliance in the HPV arm.

Unlike the ALTS trial, the detection rates of CIN2 and CIN3+ did not differ greatly between women referred with pLSIL and women referred with LSIL. The authors note that the proportion of women with pLSIL who were HPV-positive was considerably higher in this trial than in other studies conducted in the UK and US, and that many slides diagnosed as pLSIL in these countries are considered normal in Sweden. The applicability of the Swedish slide classification to the Australian setting is unknown. Although this study did not present cumulative CIN2 and CIN3+ stratified by age, it did report higher HPV positivity in women aged <35 years (pLSIL: 82.8% (78.1–86.9); LSIL: 80.5% (77.5–83.3)) than in women ≥35 years (pLSIL: 44.2% (38.4–50.1); LSIL: 68.0% (62.6–73.1)).

The number of women referred to colposcopy in the HPV arm was 1,157 (66.0%) and the number of women who actually had a colposcopy was 1,049 (59.9%), compared to 1,275 (81.4%) in the colposcopy arm. Compliance with colposcopy was higher in the HPV arm than the colposcopy arm (91% vs 81%, P<0.001).

Table 66 Cumulative diagnosis rate of CIN2–3 in women with an index cytology finding of pLSIL or LSIL managed by HPV triage or repeat cytology, Dillner et al. trial (N=3,319)[1]

Index cytology N	Cumulative risk of CIN2 -			Cumulative risk of CIN3+ -		
	HPV triage	Colposcopy	Total (test of comparison)[2]	HPV triage	Colposcopy	Total (test of comparison)[2]
pLSIL 1,335	11.4%	10.0%	10.6% (*P*=0.45)	9.1%	10.8%	10.0% (*P*=0.36)
LSIL 1,984	12.7%	9.5%	11.3% (*P*=0.031)	9.2%	9.0%	9.1% (*P*=0.94)
Total 3,319	12.2%	9.7%	11.0% (*P*=0.024)	9.2%	9.8%	9.5% (*P*=0.57)

1. Outcomes are based on linkages with cytology and pathology registers; not all women referred for colposcopy attended. 2. χ^2 test of comparison between study arms

Test accuracy (indirect evidence)

As RCT-level evidence was identified to address this research question, studies providing a lower level of evidence were excluded. However, a summary of data from the most recent systematic review reporting cross-sectional accuracy (Arbyn et al. 2013) is included in addition to data reporting on absolute and relative accuracy in the AHRQ HTA (Vesco et al. 2011).

Arbyn meta-analysis

The pooled accuracy data from the high-quality meta-analysis by Arbyn et al. (2013) are presented in Table 67 for women referred with ASCUS and in Table 68 for women referred with LSIL. At the cytology threshold of ASCUS+, the HC2 test is more sensitive than repeat cytology for women referred with either ASCUS or LSIL for the detection of CIN2+ (ASCUS referral: 90.9% [95% CI 85.7–94.4] vs 71.5% [95% CI 62.9–78.8]; LSIL referral: 96.2 [95% CI 91.4–98.3] vs 77.1% [95% CI 59.5–88.5]) and CIN3+ at an ASCUS referral threshold (94.8% [95% CI 89.6–

97.5] vs 77.9% [95% CI 64.0–87.6]) but not for LSIL referral (97.5 [95% CI 69.6–99.8] vs 84.6% [95% CI 48.6–97.0]). This difference was not statistically significant for women with LSIL referral for the detection of CIN3+ (relative sensitivity 1.15 [95% CI 0.89–1.48]).

For women referred with ASCUS, there was no difference in specificity for the detection of either CIN2+ (relative specificity 0.99 [95% CI 0.97–1.03]) or CIN3+ (relative specificity 0.99 [95% CI 0.89–1.09]). However, there was a significant reduction in specificity for women referred with LSIL for the detection of both CIN2+ (relative specificity 0.66 [95% CI 0.58–0.75]) and CIN3+ (relative specificity 0.56 [95% CI 0.37–0.84]).

Age-stratified data were only available in three studies included in the Arbyn review, and to pool this data the review authors obtained unpublished five-year age-stratified data from two of these studies. The sensitivity of HC2 triage did not vary significantly by age but the specificity always increased with age. For the detection of CIN3+ in women with LSIL referring cytology, HC2 specificity was 16.3% (95% CI 14.0–18.9) for women aged <30, 22.9% (95% CI 13.2–36.7) for women aged 30–39, 28.5% (95% CI 13.2–51.0) for women aged 40–49 and 43.4% (95% CI 22.5–67.0) for women aged 50 and over.

Table 67 Pooled accuracy from studies with HC2 triage and repeat cytology; women referred with ASCUS cytology (from Arbyn et al. 2013)

Triage test	Test cut-off	Outcome	Studies	Accuracy parameter	Pooled estimate	95% CI
HC2	RLU > (1pg/mL)	CIN2+	10	Absolute sensitivity	90.9%	85.7–94.4
HC2	RLU > (1pg/mL)	CIN2+	10	Absolute specificity	60.7%	52.9–68.0
HC2	RLU > (1pg/mL)	CIN3+	4	Absolute sensitivity	94.8%	89.6 –97.5
HC2	RLU > (1pg/mL)	CIN3+	4	Absolute specificity	56.6%	39.4–72.3
Repeat cytology	ASCUS+	CIN2+	10	Absolute sensitivity	71.5%	62.9–78.8
Repeat cytology	ASCUS+	CIN2+	10	Absolute specificity	68.4%	59.9–75.8
Repeat cytology	ASCUS+	CIN3+	4	Absolute sensitivity	77.9%	64.0–87.6
Repeat cytology	ASCUS+	CIN3+	4	Absolute specificity	57.4%	40.3–73.0
HC2 vs repeat cytology	ASCUS+	CIN2+	10	Relative sensitivity	1.27	1.16–1.39
HC2 vs repeat cytology	ASCUS+	CIN2+	10	Relative specificity	0.99	0.97–1.03
HC2 vs repeat cytology	ASCUS+	CIN3+	4	Relative sensitivity	1.14	1.06–1.22
HC2 vs repeat cytology	ASCUS+	CIN3+	4	Relative specificity	0.99	0.89–1.09

Table 68 Pooled accuracy from studies with HC2 triage and repeat cytology; women referred with LSIL cytology (from Arbyn et al. 2013)

Triage test	Test cut-off	Outcome	Studies	Accuracy parameter	Pooled estimate	95% CI
HC2	RLU > (1pg/mL)	CIN2+	6	Absolute sensitivity	96.2%	91.4–98.3
HC2	RLU > (1pg/mL)	CIN2+	6	Absolute specificity	27.7%	20.9–35.7
HC2	RLU > (1pg/mL)	CIN3+	4	Absolute sensitivity	97.5%	69.6–99.8
HC2	RLU > (1pg/mL)	CIN3+	4	Absolute specificity	24.8%	7.3–58.1
Repeat cytology	ASCUS+	CIN2+	6	Absolute sensitivity	77.1%	59.5–88.5
Repeat cytology	ASCUS+	CIN2+	6	Absolute specificity	51.2%	34.5–67.6
Repeat cytology	ASCUS+	CIN3+	4	Absolute sensitivity	84.6%	48.6–97.0
Repeat cytology	ASCUS+	CIN3+	4	Absolute specificity	44.4%	16.0–76.9
HC2 vs repeat cytology	ASCUS+	CIN2+	10	Relative sensitivity	1.23	1.06–1.43
HC2 vs repeat cytology	ASCUS+	CIN2+	10	Relative specificity	0.66	0.58–0.75
HC2 vs repeat cytology	ASCUS+	CIN3+	4	Relative sensitivity	1.15	0.89–1.48
HC2 vs repeat cytology	ASCUS+	CIN3+	4	Relative specificity	0.56	0.37–0.84

AHRQ HTA

The AHRQ report (Vesco et al. 2011) included four cohort studies reporting on test accuracy and pooled data from three of these for women referred with ASCUS cytology for the detection of CIN2+. The pooled difference in sensitivity between HC2 triage and repeat cytology was 12% (95% CI 0.2–23.9) and no difference in specificity was observed (p=0.65 for the combined difference). This is in agreement with the results of the Arbyn et al. (2013) review for this cytology referral threshold.

Prognostic value of test results

The section 'Prognostic value of test results' in the chapter on HPV primary testing (Page 171) reports on the predictive value of test results for the development of CIN2+ and CIN3+ in longitudinal studies (HPV and cytology) from two large cohort studies and can be informative in a consideration of HPV triage of pLSIL and dLSIL, although outcomes for women with positive cytology include all women with pLSIL+ cytology and is therefore a broader group than is considered here. Considering the consistency of the body of evidence discussed above these prognostic data are considered of limited additional value in addressing this research question.

Single HPV test to exit the screening program

Secondary research question 2.3

What is the comparative safety, effectiveness and cost-effectiveness of using one HPV test for women exiting the program at age 65 years and over, compared with the existing protocol?

Summary 8: Summary of effectiveness—single HPV exit test

No studies were identified that compared the use of a single HPV test to that of two consecutive cytology tests as an exit strategy.

A case series of 56 women with invasive cervical cancer (Dinkelspiel et al. 2012) found that the majority of cervical cancer cases in women aged 65 and over were in women not meeting the criteria for an exit test.

In the follow-up of the ARTISTIC trial, significantly fewer women of any age with a negative HPV test develop high-grade precancerous cervical lesions than those testing negative on a single cytology test. Data on the prognostic value of two consecutive negative cytology tests were not provided.

A modelled analysis of screening, diagnosis and treatment in the Australian setting is required to explore the benefits and harms of the alternative exit testing strategies.

Conclusion

- The majority of cervical cancer cases in women aged 65 and over are in women not meeting the criteria for an exit test.

Table 69 Inclusion criteria for identification of studies relevant to an assessment of effectiveness of intervention

Selection criteria	Inclusion criteria
Population	Women ≥65 years undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Intervention	HPV testing with or without cytology as either a co-test or a reflex test
Comparator(s)	Cytology testing
Outcomes	Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC) Cervical cancer incidence Mortality
Search period	2008–2012 (HTAs); 2011–2012 (primary studies)
Language	English
Study design	Any

The most recent systematic review or HTA identified and included for HPV exit testing is listed in Table 70. The inclusion criteria used are listed in Table 69.

Table 70 Relevant systematic reviews and meta-analyses identified for effectiveness of a single HPV test to exit the screening program

Study identified in review	Study title	Studies included in this review
ACS Saslow et al. 2012	Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam S, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS, Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER, ACS-ASCCP-ASCP Cervical Cancer Guidelines Committee 2012. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer	Nil

Saslow et al. (2012) conducted an HTA to support an update to the American Cancer Society guidelines. This review did not include a comparison of alternative testing strategies for exiting screening programs. Neither was the evidence for the effectiveness of a single HPV test in comparison to no screening in women over age 65 assessed. This HTA reviewed the evidence for comparisons of different testing strategies to no screening in women over 65 years (cytology alone or HPV and cytology co-testing every three or five years was considered). No clinical studies were identified that addressed these questions. Only a single modelling study that does not address the research question for this review was identified (it considered no screening in comparison to continued cytology screening).

Saslow et al. (2012) comment that "screening [in women over 65 years] would detect a very small number of new cases of CIN2+ and prevent very few cervical cancers and even fewer cancer deaths. The extended natural history of the disease also makes it less likely that newly detected CIN3 will have time to progress to invasive cancer in the woman's lifetime. There is also evidence that screening is associated with potential harms, including anxiety and discomfort during cytology sampling of some older women due to vaginal atrophy and cervical stenosis. The choice of exact age at which to cease screening is arbitrary." Further, they state that "In well screened older women, HSIL rates are low (Castle et al. 2009) and cervical cancer is rare (Saslow et al. 2002) Most new cases of cervical cancer in U.S. women >65 years are in unscreened or infrequently screened women [Mandelblatt et al. 1986; Sawaya et al. 2000]. Reducing the burden of cervical cancer on older women is likely best achieved by focusing on screening those who have not been adequately screened."

These guidelines recommended that "women over 65 years of age with evidence of adequate negative prior screening and no history of CIN2+ or cervical cancer within the last 20 years should not be screened for cervical cancer with any modality (weak recommendation), adequate negative prior screening is defined as 3 consecutive negative cytology results or 2 consecutive negative cotests within the last 10 years before ceasing screening, with the most recent test being within the past 5 years."

The AHRQ HTA also does not specifically address the question of exit testing strategies, but it does include an informative discussion on the negative predictive value (NPV) of HPV testing. Their conclusions are that "women screening negative on HPV testing have a nearly identical reduced long-term risk of developing CIN3+ as women that are HPV negative/cytology negative. The high NPV associated with HPV negative testing alone, particularly in older women, might inform extended screening intervals for such women, with no need for cytology testing at all in HPV negative women."

Included studies

No studies comparing the effectiveness of a single HPV test with cytology testing as an exit test were identified in the update searches.

One retrospective case series (Dinkelspiel et al. 2012) was identified which examined the screening history of cancer cases in women aged over 65 to inform considerations of the safety of exiting screening in older women. In the absence of studies providing direct evidence of a higher level, this study was included for review. This study was also included in the assessment of evidence for the age range of screening (from page 70).

Cohort studies providing data on the predictive value of test results that were not conducted with the objective of assessing exit testing strategies were not included. However, one RCT (the ARTISTIC trial, Kitchener et al. 2011a) included in the review of evidence for HPV testing as a primary screening test provided data on cumulative detection of CIN2+ and CIN3+ after three screening rounds by baseline testing status in women aged 50 to 64 years. These data provide some information on the prognostic value of different testing strategies for the development of high-grade precancerous cervical lesions. This study was the only included RCT of HPV testing with follow-up over three screening rounds and that provides data for an age category close to 65 years. In the absence of evidence from studies designed to compare different exit-testing strategies, these data were considered in the evidence for the effectiveness of exit-testing strategies.

A master list of the primary studies included in the review of HPV exit testing is provided in Table 71.

Table 71 Master list of trials included for evaluation of effectiveness of a single HPV test to exit the screening program

Study identified in review	Study title	Comments
Dinkelspiel et al. 2012	Dinkelspiel H, Fetterman B, Poitras N, Kinney W, Cox JT, Lorey T, & Castle PE (2012). Screening history preceding a diagnosis of cervical cancer in women age 65 and older. Gynecol Oncol 126, 203–206	Identified in updated systematic review
ARTISTIC	Kitchener HC, Gilham C, Sargent A, Bailey A, Albrow R, Roberts C, Desai M, Mather J, Turner A, Moss S, Peto J (2011). 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 47, 864–871	Identified in updated systematic review

Study characteristics, quality and applicability

Dinkelspiel et al. (2012) was a retrospective case series selecting cancer cases from the laboratory databases of North Carolina Kaiser Permanente (US). The characteristics of this study are presented in Table 72. This study does not provide comparative evidence for the different exit testing strategies (NHMRC level IV evidence). The study reported the testing history of these cases over an era where there were two alternative exit-testing strategies (2003 to 2008), either three consecutive conventional cytology smears or a single negative conventional and HPV co-test. The proportion of women receiving co-testing over this era was not reported. The overall rate of cancer incidence was compared to that in a historical era where three consecutive conventional cytology tests was the only exit strategy. However, there is a high risk of bias for these estimates of cancer incidence rates because the investigators relied on laboratory test use to estimate person-time at risk. A single HPV test without cytology was not used as an exit strategy; hence the results of this study are not directly applicable to the research question. The characteristics of this study are also discussed above (see Table 25, page 71).

Table 72 Characteristics and quality of included study for HPV exit testing

Author, year Setting	Population (N)	Test comparison/s	Study design	Source Quality and applicability
Dinkelspiel et al. 2012 USA Kaiser Permanente North Carolina Number of sites NR Cases diagnosed 2003–2008	N=56 cases Inclusion criteria • Women aged ≥65 yrs • Health insurance members • Screened & unscreened women Exclusion criteria • NR Patient characteristics • Mean age 75 (median 73, range 65–101)	Index test Exit testing strategies: Single negative HPV co-test or 3 consecutive negative conventional cytology tests (alternative strategies 2003–2008) Comparator test Nil	Study design: Case series Test threshold: NR Outcomes Proportion of cancer cases meeting cytology exit test criteria, proportion of women meeting co-testing exit criteria diagnosed with cancer (case series data) Incidence of cervical cancer, proportion of cancer cases not meeting stopping criteria (cohort data with historical comparison)	Systematic review search Screening study NHMRC level IV Quality: High risk of bias Applicability: low No comparative data Intervention HPV co-testing rather than HPV testing alone

Abbreviations: CC = conventional cytology; LBC = liquid-based cytology; NR = not reported

The characteristics of the ARTISTIC trial are discussed in detail below, under consideration of evidence for primary HPV testing (see Table 81, page 149). The recent publication of cumulative CIN2+ and CIN3+ detection over three screening rounds, stratified by age provides some information on the predictive value of baseline test results. This provides evidence at a cohort level (level III-2) for the effectiveness of screening interventions. The outcomes are lower than those considered in the hierarchy suggested by Arbyn et al. 2009 (see 'Benefits of screening', page 22), but are presented in the absence of other comparative evidence.

Incidence and mortality from cervical cancer (direct evidence)

No studies assessing the impact of a single HPV exit test on incidence or mortality rates of invasive cervical cancer compared to the current strategy of two normal conventional cytology tests within the past five years were identified.

The Dinkelspiel et al. (2012) Kaiser Permanente study reported that the majority of women diagnosed with invasive cervical cancer over the age of 65 had not met the criteria for exiting screening (75% over the period 2003–2008 with HPV co-testing as an alternative exit strategy; 91% in the era where only three consecutive negative conventional cytology tests was the exit strategy). Three of 46,401 women aged 65 or over with one or more negative co-tests were diagnosed with cervical cancer (2.3/100,000/year). These women also met the criteria of three consecutive negative cytology tests. Of all women aged 65 and over, 4.2/100,000/year were diagnosed with cervical cancer in the same era, or 4.0/100,000/year during the period 1988–1994 when only three consecutive conventional cytology tests was used as an exit strategy. However as the majority of cases did not meet the exit criteria, a direct comparison to the rate in women with a negative co-test is not considered informative. Also, these incidence rates are considered to have a high risk of bias.

Detection of precancerous cervical lesions

No studies comparing the use of a single HPV test as an exit strategy with alternative testing were identified in the primary HPV search.

Prognostic value of test results

The prognostic value of HPV and cytology tests is discussed in the section 'Prognostic value of test results' on page 171. Two cohort studies were included (Kitchener et al. 2011b, Katki et al. 2011) neither of which have all data stratified by age. They both report a low risk of developing CIN2+ or CIN3+ over five years following a HPV negative test, this risk is lower than the risk following a negative cytology test and similar to the risk of both tests being negative. Discussion of these studies is supplementary to the systematic review.

In the Kitchener study (2011b) which is the extended follow-up from the ARTISTIC trial, the predictive value of different baseline HPV results in women aged 50 to 64 for the cumulative detection of CIN2+ and CIN3+ after three screening rounds is reported. The rates of cumulative CIN detection following one or two negative cytology tests in women aged 50 to 64 years were not reported.

A HPV negative test in women 50 to 64 years had a cumulative CIN2+ detection rate over six years of 0.16% (95% CI 0.07 –0.34 and a cumulative CIN3+ detection rate of 0.05%, 95%CI 0.01–0.19; Table 73). A rate of CIN2+ of 0.16% over six years equates to 1.6 women per 1,000

women screened. The detection rate in the whole study population (age 20–64) who tested HPV negative was 0.87% (95% CI 0.70–1.06) for CIN2+ and 0.28% for CIN3+ (0.18–0.40).

Table 73 Six-year cumulative CIN2+/CIN3+ rates associated with a negative HPV baseline test result, ARTISTIC trial (Kitchener et al. 2011b)

Age	Cumulative CIN2+ (95% CI)	Cumulative CIN3+ (95% CI)
20–64	0.87% (0.70%–1.06%)	0.28% (0.18%–0.40%)
20–24	3.04% (1.94%–4.45%)	0.57% (0.16%–1.45%)
25–34	1.44% (1.03%–1.97%)	0.53% (0.29%–0.88%)
35–49	0.48% (0.32%–0.71%)	0.18% (0.08%–0.34%)
50–64	0.16% (0.07%–0.34%)	0.05% (0.01%–0.19%)

The Katki cohort (2011) found that the low risk of CIN3+ in women negative by HPV and Pap testing at enrolment was further reduced for women with no history of a previous abnormal Pap test (hazard ratio [HR] 0·21, 95% CI 0·13–0·49; p=0·0002) and for women aged 50 years and older compared with women aged 30–34 years (HR 0·22, 0·11–0·32; P<0·0001).

Discussion

When assessing the chance of diagnosis of cervical cancer in women who have exited the screening program, the trade-off in terms of the harms of screening (false positive rates etc.) must also be considered. This will be explored in the modelled analysis. It is noteworthy that Dinkelspiel et al. (2012), in accord with a discussion by Saslow et al. (2012), found that the majority of cervical cancer cases in women were in those not meeting the criteria for an exit test.

Immediate versus delayed colposcopy for pLSIL/LSIL HPV positive screenees

This question is addressed in the section 'Results of assessment – effectiveness: HPV testing (scenario 3)', page 174.

Results of assessment – effectiveness: HPV testing (scenario 3)

Primary HPV screening alone or in combination with LBC

Primary research question 3

What is the comparative safety, effectiveness and cost-effectiveness of HPV testing as the primary screening test in women aged 25 to 65 years every five years, compared with the protocol used in the current Australian cervical screening program?

Secondary research question 3.1

What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a co-test in the above scenario, compared with the protocol proposed in the primary question?

Secondary research question 3.2

What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a reflex test to triage women with positive HPV test results, compared with the protocol proposed in the primary question?

Summary 9: Summary of effectiveness—RCTs of HPV screening strategies

One cluster randomised trial conducted in India (Sankaranarayanan 2009) demonstrates that a single round of HC2 testing can significantly reduce both the cervical cancer incidence and mortality in comparison to a single cytology screen in a previously unscreened population in a low-resource setting.

Evidence from seven RCTs (conducted in countries with established screening programs) that considered various HPV screening strategies, has been assessed in order to provide evidence for the questions posed for this review. All were regarded to be fair quality and were conducted in unvaccinated women. The studies assessed are as follows:

- One RCT (N =49 196) was of HPV primary screening.
- Two RCTs (N = 150,842) were of HPV screening with cytology triage.
- The remaining four RCTs (N =128,149) were of HPV and cytology co-testing.
- The HPV test used in the majority of studies was HC2; however, two studies used a polymerase chain reaction (PCR) test.
- Five studies provided a comparison with conventional cytology, one with LBC and one LBC with HPV triage.
- The studies varied in their characteristics, including the threshold for referral to colposcopy, follow-up testing procedures and included age ranges.
- No studies recruited women aged over 65 years.
- Five studies reported results over two rounds; the protocol for the second round of screening differed from the first round in three of these.

Trials were not powered to detect a difference in invasive cervical cancer incidence. There was a broad trend of a reduction in the number of cervical cancer cases detected in the second round in the intervention arm. This was significant in one trial of HPV and cytology co-testing (POBASCAM; RR: 0.29, 0.10–0.87; p=0.031).

Cumulative detection ratios for CIN3+ did not differ significantly between arms in four of five trials in the older age group (≥30 or 35 years); one study comparing HPV primary testing alone to conventional cytology (NTCC Phase II) had a significantly increased cumulative detection of CIN3+ over two screening rounds.

In the first screening round, three of seven trials demonstrated a significantly greater detection of CIN3+ in women aged ≥30–35 years with a HPV screening strategy compared with a cytology-only screening strategy. This was also observed in two of four trials reporting data for women aged ≤30–35 years. More trials reported increased detection of CIN2+ in the first screening round; five of seven trials in women ≥30–35 years and three of four in women aged ≤30–35 years. The extent to which additional detected precancerous lesions would progress to clinically significant disease is uncertain.

For five trials reporting data for an older age group (≥30–35 years), a consistent reduction in the relative detection ratio of CIN3+ was seen in a second screening round and was statistically significant in three of four, despite differences in referral thresholds and follow-up procedures. A similar trend was observed for the detection of CIN2+; there was a significant reduction in three of five trials.

The colposcopy referral rates for HPV-based screening strategies were increased in five of six studies in women aged ≥30–35 years, and in all five studies in women aged ≤30–35 years.

The overall body of evidence for the effectiveness of HPV based screening strategies (with or without triage or co-testing) is considered good consisting of several large RCTs (level II evidence) across multiple rounds whose findings are relatively consistent for the surrogate outcomes reported despite differences in design. The populations considered in these trials are similar to the target population and the results are likely to be applicable to unvaccinated women in the Australian context with some caveats.

In addition to the systematically reviewed RCTs of HPV-based screening, supportive data from two large cohort studies (Katki et al. 2011, Kitchener et al. 2011b) suggest that longer screening intervals up to 5 years may be appropriate for women who are HPV negative due to the high negative predictive value of HPV testing.

Increased detection of precancerous lesions in a first screening round followed by decreased detection in the second round may indicate a benefit due to increased detection and treatment of progressive lesions, or simply earlier detection. Any potential harm in terms of increased colposcopy rates and the treatment of additional non-progressive lesions should also be considered. It appears that the balance of benefits and potential harms is more favourable in women ≥30–35 years. A modelled analysis considering the potential long-term benefits and trade-offs of these screening strategies in the Australian setting has been conducted.

Conclusions

- HPV testing has a high negative predictive value (the probability that a negative test result is a true negative).

- The balance of benefits from increased detection of precancerous cervical lesions and potential harms from increased colposcopy referrals is more favourable for women aged ≥30–35 years for all HPV based screening strategies.

Table 74 Inclusion criteria for identification of studies relevant to the assessment of the effectiveness of primary HPV screening

Selection criteria	Inclusion criteria
Population	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Intervention	HPV testing alone (1º question) HPV testing with LBC as either a co-test or a reflex test (2º questions)
Comparator(s)	Conventional cytology (1º question) HPV as a primary test (2º questions)
Outcomes	Cervical cancer incidence Mortality Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC) in women with a possible or definite HSIL cytology result
Reference standard	Colposcopy with biopsy for positives
Search period	2008–2012 (HTAs); 2011–2012 (primary studies)
Language	English
Study design [a]	RCT

Further detail provided in Box 2, page 39.
Abbreviations: AIS = adenocarcinoma in situ; HSIL = high-grade squamous intraepithelial lesion; LBC = liquid-based cytology; SCC = squamous cell carcinoma; HPV = human papillomavirus; HTAs = health technology assessments
[a] Post-hoc inclusion criteria based upon the body of evidence identified

The most recent systematic reviews and HTAs identified and included for the assessment of evidence for HPV testing as a primary cervical screening strategy are listed in Table 75. The inclusion criteria used are listed in Table 74.

Table 75 Relevant systematic reviews and meta-analyses included for effectiveness of HPV-based strategies for primary cervical cancer screening

Study identified in review	Study title	Studies included in this review
Peirson et al. 2012	Peirson, L., Fitzpatrick-Lewis, Ciliska, D and Warren, R (2012). Screening for cervical cancer. Canadian Cervical Screening Initiative	Sankaranarayanan et al. 2009 (India) Finnish study
AHRQ HTA Vesco et al. 2011	Vesco KK, Whitlock EP, Eder M, Lin J, Burda BU, Senger CA, Holmes RS, Fu R, Zuber S. Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 86. AHRQ Publication No. 11-05156-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; May 2011 Whitlock EP, Vesco KK, Eder M, Lin JS, Senger CA & Burda BU (2011). Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 155, 687–215	NTCC Phase I & II Finnish study Swedscreen ARTISTIC POBASCAM
Arbyn et al. 2012	Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G, Koliopoulos G, Naucler P, Sankaranarayanan R, Peto J (2012) Evidence regarding Human Papillomavirus testing in secondary prevention of cervical cancer. Vaccine 30S, F88–F99	Meta-analysis of 49 studies

Incidence and mortality from cervical cancer (direct evidence)

Included studies

The most recent high-quality systematic review of HPV screening compared to conventional cytology identified two RCTs eligible for inclusion based on the outcomes of cervical cancer incidence and mortality (Peirson et al. 2012). One of these RCTs (Sankaranarayanan et al. 2009) was conducted in an unscreened population in India. The other RCT (Anttila et al. 2010) was conducted in Finland and although it did report cervical cancer incidence, the low number of events and wide confidence intervals do not make this a meaningful outcome. This trial is therefore considered in the section on the detection of precancerous cervical lesions (indirect evidence) and only the Indian trial data are presented in this section.

Table 76 Master list of trials included for evaluation of effectiveness of HPV versus cytology for primary screening with outcomes of incidence and mortality

Study identified in review	Study title	Comments
Sankaranarayanan et al. 2009	Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, Hingmire S, Malvi SG, Thorat R, Kothari A, Chinoy R, Kelkar R, Kane S, Desai S, Keskar VR, Rajeshwarkar R, Panse N, Dinshaw KA. HPV screening for cervical cancer in rural India. N Engl J Med. 2009; 360 (14):1385–94	Included in Peirson et al. 2012 Included in Vesco et al. 2011

Study characteristics, quality and applicability

The review conducted for the Canadian Taskforce on Preventive Health Care was considered a high-quality systematic review (Peirson et al. 2012). The inclusion criteria were that studies must report incidence or mortality of invasive cervical cancer; both RCTs and observational studies with comparison groups were included. The characteristics of this review are summarised in Table 77.

Table 77 Characteristics of most recent systematic reviews and HTAs of HPV screening

Author, year Country	Objective and methods	Inclusion/exclusion Included studies	Quality and applicability Conclusion
Peirson et al. 2012 Canada	Objective: Does either primary or reflex HPV testing reduce incidence or mortality from invasive cervical cancer compared to conventional cytology? (Qu 1b) Literature review: • Medline, EMBASE, Cochrane • 1995 to February 2011	Inclusion/exclusion criteria: *Study design:* Meta-analyses, RCTs, observational studies *Intervention:* HPV testing *Comparator:* Conventional Pap *Outcomes:* Mortality, cervical ca incidence *Language:* English or French 2 studies	Quality: High Applicability: High Conclusion An RCT in India showed that even a single lifetime HPV test significantly decreased incidence of and mortality from invasive cervical cancer compared to no screening.

Abbreviations: ca = cancer; CC = conventional cytology; LBC = liquid-based cytology; NR = not reported; Qu = question

The Indian trial (Sankaranarayanan et al. 2009) included in the Peirson et al. (2012) review was a cluster randomised trial which compared HPV testing (HC2), conventional cytology and visual inspection of the cervix with acetic acid (VIA) to a never-screened control group. The study was conducted in rural India, on a population of never-screened women aged 30–59 years old and is of limited applicability to the Australian population where most women are screened for cervical cancer. The incidence of disease is lower, the quality of cytological assessment is likely to be higher and standards of follow-up care are likely to differ to those in Australia. The Canadian review rated the Indian trial as of moderate quality according to GRADE criteria (Guyatt et al. 2008). The characteristics of this study are summarised in Table 78.

Table 78 Characteristics and quality of studies reporting on the incidence and mortality from cervical cancer of HPV screening compared to cytology

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Sankaranarayanan et al. 2009 India 497 villages in 52 clusters, assigned to 4 groups of 13 clusters each Unscreened population	N=131,806 Inclusion criteria: Age 30–59 yrs 'Healthy' Currently or previously married Intact uterus Lives in study cluster Exclusion criteria Pregnant Uterine prolapsed History of cervical cancer Patient characteristics Mean age: HC2: 39 Cytology: 39 Control 40	HC2 vs conventional cytology vs VIA vs control	Study design: Cluster randomised trial Test threshold: HC2:≥ 1 pg/mL Criteria for colposcopy referral: Positive screening test Outcomes Cumulative data over 8 years for cancer outcomes.	Quality: Moderate (Peirson GRADE rating)/Fair (AHRQ rating) Applicability: Low

Abbreviations: AHRQ = Agency for Healthcare Research and Quality; GRADE = Grading of Recommendations Assessment, Development and Evaluation; HC2 = Hybrid Capture 2; VIA = visual inspection with acetic acid

Incidence and mortality

Compared to the unscreened control group, the HPV-screened group had a significantly reduced number of cancer deaths (adjusted HR, 0.52 [95% CI, 0.33–0.83]) and incidence of Stage II or higher cervical cancer (adjusted HR, 0.47 [95% CI, 0.32–0.69]) (Vesco et al. 2011, page 24). A single round of screening with conventional cytology did not significantly reduce the numbers of Stage II or higher cervical cancers or deaths, compared to the unscreened control group. The study demonstrates that a single round of HC2 testing can significantly reduce both the cervical cancer incidence and mortality in a previously unscreened population in a low-resource setting.

Detection of precancerous cervical lesions (indirect evidence)

The most recent high-quality systematic review of HPV-based screening and conventional cytology only included studies which reported on incidence and mortality from cervical cancer (Peirson et al. 2012). The present review also relies on evidence about the detection of precancerous cervical lesions on the basis that cervical cancer is rare in screened populations and therefore trials designed to detect differences in incidence and mortality of cervical cancer in these populations are unfeasible. Reduced CIN3+ is considered as an acceptable proxy outcome of trials evaluating new preventative strategies (Arbyn et al. 2009).

Thus the current review considers and updates evidence from the AHRQ HTA (Vesco et al. 2011) which considered evidence on the detection of CIN3+ and the relative accuracy of HPV testing and conventional cytology. A systematic literature search for HPV primary studies was conducted to identify studies that were published since the AHRQ systematic literature search, including studies from August 2011 to November 2012 to update this report. Results are presented for women aged 30 or 35 years and older and for women aged younger than age 30 or 35 years for consistency with the AHRQ report and due to the known decrease in HPV prevalence with age (Vesco et al. 2011).

LBC is considered to provide no significant difference in sensitivity for the detection of CIN2+ compared to conventional cytology (see 'Results of Assessment— Effectiveness: Liquid-based cytology (scenario 2)', page 90). Therefore, studies of LBC or conventional cytology were pooled and the body of evidence considered as one.

Included studies

The AHRQ HTA conducted for the US Preventive Services Task Force (USPSTF) (Vesco et al. 2011) was considered a high-quality systematic review and included studies providing data on the detection of CIN3+ over multiple rounds of screening. The characteristics of this review are summarised in Table 61.

For studies evaluating HPV testing, the AHRQ included studies that provided evidence on the absolute and relative test performance, applied a reference standard of colposcopy and/or biopsy to all those screening positive and at least a random sample of those screening negative, and were conducted in routine screening populations. The high-quality AHRQ review considered evidence to September 2010 and conducted an ancillary search which identified additional studies published to August 2011. Six RCTs were identified and included that reported relative test performance data, thus the measures of test performance were valid although the reference standard was not applied to all subjects.

Table 79 Characteristics of most recent systematic reviews and HTAs of HPV screening

Author, year Country	Objective and methods	Inclusion/exclusion Included studies	Quality and applicability Conclusion
Vesco et al. 2011 AHRQ USA	Objective: What are the benefits of using HPV testing as a screening test, either alone or in combination with cytology, compared to conventional cervical cytology (KQ3)? Literature review: • DARE, HTA database, Cochrane Systematic Reviews & Trials Registry, PubMed, MEDLINE, PsychINFO, contacted authors • January 2000–September 2010 • Ancillary search 1 September 2010–3 August 2011	Inclusion/exclusion criteria: *Study design:* Systematic ref std, ref std only in positives for absolute accuracy, exclude case-control *Intervention:* HPV (HC2 or PCR) alone or in combination with cytology *Comparator:* Conventional or liquid-based cytology *Outcomes:* Absolute & relative test performance for detection CIN2+ *Exclusion:* Poor quality studies, non-English language 22 unique studies in 48 publications • Primary HPV alone: 2 RCTs, 6 observational studies • Reflex cytology: 1 RCT • Co-testing: 4 RCTs, 4 cohort studies • Reflex HPV: 2 RCTs, 4 cohort studies	Quality: High Applicability: High Conclusion One-time HPV screening was more sensitive than cytology for detecting CIN3+/CIN2+ but was less specific. Primary HPV detected more cases of CIN3+ or cancer in women older than 30 years. Studies showed mixed results of co-testing in women aged 30 years or older compared with cytology alone, with no clear advantage over primary HPV screening. Incomplete reporting of results for all screening rounds, including detection of disease and colposcopies, limits our ability to determine the net benefit of HPV-enhanced testing strategies.
Arbyn et al. 2012	Objective: To update and extend previously conducted meta-analyses and systematic reviews, on the performance of HPV testing as primary cervical screening test Literature review: • Specialised Register of the Cochrane Gynaecological Cancer Group, Non-trials database of the Cochrane Gynaecological Cancer Group, CENTRAL on the Cochrane Library, The Health Technology Assessment Database, MEDLINE, EMBASE, CINAHL, LILACS) • Dates NR	Inclusion/exclusion criteria: *Study design:* Cross-sectional studies or RCTs with either all verified with ref standard, only test positives verified with ref standard or random sample of negatives verified with ref standard. *Intervention:* HPV (any test) in women representative of the general population *Comparator:* Conventional or liquid- based cytology *Outcomes:* Absolute & relative test performance for detection CIN2+/CIN3+ 49 (cross-sectional accuracy), 4 (longitudinal outcomes of RCTs), 7 (longitudinal outcomes of cohort studies)	Quality: High Applicability: High Conclusion High-risk HPV testing with a clinically validated assay is more effective than cervical cytology for women 30 years and older; screening intervals for HPV-negative women can be safely extended to at least 5 years.

Abbreviations: ca = cancer; CC = conventional cytology; LBC = liquid-based cytology; NR = not reported; Qu = question; ref = reference

The AHRQ HTA (Vesco et al. 2011) included six RCTs of HPV-based primary screening strategies compared to cytology that provided data on the detection of precancerous cervical lesions. As RCT level evidence was identified to address this research question, studies providing a lower level of evidence were excluded.

The AHRQ also included cohort studies reporting absolute accuracy. As these studies provided a lower level of evidence than the RCTs described above, they are not included in the current review and the update search did not retrieve studies of this design. However, results from the RCTs are

supplemented with accuracy data from these cohort studies and the meta-analyses by Arbyn et al. (2012).

The meta-analysis by Arbyn et al. (2012) was considered a high-quality review and updates previous publications by the same group (Koliopoulous et al. 2007; Arbyn et al. 2006). The review included non-randomised controlled trials using HPV testing and concomitant cytological screening on the same woman and randomised controlled trials comparing cytological screening versus HPV testing. The studies were required to be in women representative of the general population and the reference standard, in the form of colposcopy and/or biopsy was required to be applied on at least all the Pap test or HPV test positives and the outcomes considered were the presence of CIN2+ and CIN3+. The review included 49 studies of cross-sectional accuracy of which eight were randomised trials.

One additional RCT (HPV FOCAL, Ogilvie et al. 2010) was excluded from the AHRQ review as it did not contain any relevant outcomes. A new publication (Ogilvie et al. 2012) reporting results for this trial was identified in the update search and this trial is included in the current review. In addition, new studies updating results for three of the six RCTs included in the AHRQ were identified in the update search; these new publications are identified in Table 80.

Table 80 Master list of trials included for evaluation of effectiveness of HPV versus cytology for primary screening with outcomes of detection of precancerous cervical lesions

Study identified in review	Study publications (publications identified in the update search in bold)	Comments
NTCC phase II	Ronco G, Brezzi S, Carozzi F et al. The New Technologies for Cervical Cancer Screening randomised controlled trial: an overview of results during the first phase of recruitment. Gynecol Oncol. 2007;107(Suppl 1):S230–232 Ronco G, Cuzick J, Segnan N et al. HPV triage for low grade (L-SIL) cytology is appropriate for women over 35 in mass cervical cancer screening using liquid based cytology. Eur J Cancer. 2007;43:476–480 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 Oncology. 2010;11:249–257 Ronco G, Giorgi-Rossi P, Carozzi F et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst.* 2008;100:492–501 Ronco G, Segnan N, Giorgi-Rossi P et al. Human papillomavirus testing and liquid-based cytology: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. J Natl Cancer Inst. 2006;98:765–774	Included in AHRQ
Finnish trial	Anttila A, Hakama M, Kotaniemi-Talonen L et al. Alternative technologies in cervical cancer screening: a randomised evaluation trial. *BMC Public Health.* 2006;6:252 Anttila A, Kotaniemi TL, 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. 2010;340:c1804 Kotaniemi-Talonen L, Anttila A, Malila N et al. Screening with a primary human papillomavirus test does not increase detection of cervical cancer and intraepithelial neoplasia 3. *Eur J Cancer.* 2008;44:565–571 Kotaniemi-Talonen L, Nieminen P, Anttila A et al. Routine cervical screening with primary HPV testing and cytology triage protocol in a randomised setting. Br J Cancer. 2005;93:862–867 Leinonen M, Nieminen P, Kotaniemi-Talonen L et al. Age-specific evaluation of primary human papillomavirus screening vs conventional cytology in a randomized setting. J Natl Cancer Inst. 2009;101:1612–1623 **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. 2012 (in press)** **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. Int J Cancer. 2012 (in press)**	Included in AHRQ New publications identified in updated literature search
NTCC Phase I	Ronco G, Brezzi S, Carozzi F et al. The New Technologies for Cervical Cancer Screening randomised controlled trial: an overview of results during the first phase of recruitment. Gynecol Oncol. 2007;107(Suppl 1):S230–232 Ronco G, Cuzick J, Segnan N et al. HPV triage for low grade (L-SIL) cytology is appropriate for women over 35 in mass cervical cancer screening using liquid based cytology. Eur J Cancer. 2007;43:476–480	Included in AHRQ New publication identified in updated literature search

Study identified in review	Study publications (publications identified in the update search in bold)	Comments
	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 Oncology. 2010;11:249–257 Ronco G, Segnan N, Giorgi-Rossi P et al. Human papillomavirus testing and liquid-based cytology: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. J Natl Cancer Inst. 2006;98:765–774 **Rossi PG, Carozzi F, Collina G et al. HPV testing is an efficient management choice for women with inadequate liquid-based cytology in cervical cancer screening. Anatomic Pathology. 2012;138:65–71**	
POBASCAM	Bulkmans NW, Berkhof J, Rozendaal L et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet.* 2007;370:1764–1772 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;110:94–101 **Rijkaart DC, Berkjof J, Rozendaal L et al. Human Papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial**	Included in AHRQ New publication identified in updated literature search
Swedscreen	Elfgren K, Rylander E, Rådberg T et al. Colposcopic and histopathologic evaluation of women participating in population-based screening for human papillomavirus deoxyribonucleic acid persistence. Am J Obstet Gynecol. 2005;193:650–657 Naucler P, Ryd W, Tornberg S et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. J Natl Cancer Inst. 2009;101:88–99 Naucler P, Ryd W, Tornberg S et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. New Engl J Med. 2007;357:1589–1597	Included in AHRQ
ARTISTIC	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;13:1–150 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 Oncology. 2009;10:672–682 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;95:56–61 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 Journal of Cancer. 2011;47:864–871 Sargent A, Bailey A, Almonte M et al. Prevalence of type-specific HPV infection by age and grade of cervical cytology: data from the ARTISTIC trial. Br J Cancer. 2008;98:1704–1709 Sargent A, Bailey A, Turner A et al. Optimal threshold for a positive Hybrid Capture 2 test for detection of human papillomavirus: data from the ARTISTIC trial. J Clin Microbiol. 2010;48:554–558	Included in AHRQ
HPV FOCAL	Ogilvie GS, van Niekerk DJ, Krajden M et al. A randomized controlled trial of	Excluded from AHRQ

Study identified in review	Study publications (publications identified in the update search in bold)	Comments
	Human Papillomavirus (HPV) testing for cervical cancer screening: trial design and preliminary results (HPV FOCAL trial). BMC Cancer. 2010;10:111 **Ogilvie GS, Krajden M, van Niekerk DJ et al. Primary cervical cancer screening with HPV testing compared with liquid-based cytology: results of round 1 of a randomised controlled trial – the HPV FOCAL study. Br J Cancer. 2012:107:1917–1924**	New publication identified in updated literature search

Study characteristics, quality and applicability

None of the seven included trials was conducted in HPV-vaccinated populations. Therefore, the prevalence and incidence of cervical cancer and precancerous cervical lesions will differ to that expected in Australia in the future, particularly in young women.

Of the seven included studies, five used conventional cytology in the comparator arm, one used LBC and one used LBC with HPV triage. For the intervention arm, one study used HPV primary screening, two used HPV screening with cytology triage (one conventional cytology and one LBC) and the remaining four used HPV and cytology co-testing (two with conventional cytology and two with LBC). The HPV test used in the majority of studies was HC2; however, two studies used a PCR test.

Thresholds for referral to colposcopy differed across studies and across sites in some studies. Similarly, each study had different follow-up testing procedures for abnormal results which were not referred directly to colposcopy. Two studies only recruited women aged more than 30 years and no studies recruited women aged over 65 years. The Swedescreen study had a narrow recruitment age of 32–38 years.

Five studies reported results over two rounds; the protocol for the second round of screening differed from the first round in three of these. Increased detection of precancerous cervical lesions in round one may represent additional diagnosis of prevalent progressive and/or regressive disease, or earlier diagnosis (Vesco et al, 2011, page 41). Having the same protocol in the second round for both arms of the trial ensures greater comparability of endpoint assessment. Where both arms undergo HPV testing at the second round all lesions which are detectable by HPV screening will be identified at the second round and any decrease in detection rates between the arms should represent the detection of only incident, rather than incident and prevalent disease, plus lesions which would have otherwise regressed. If the second screening round relies on cytology then the difference between arms will represent both lesions which are detected earlier by HPV than cytology and lesions which would have otherwise regressed. Therefore, further screening rounds may be required to fully assess whether an excess of lesions which would otherwise regress are being detected; however, a very large difference (eg a ratio of >2.5) is considered strongly suggestive of excess detection of regressive lesions (Ronco & Segnan 2007; Ronco et al. 2012).

All studies were rated as fair quality by the AHRQ and are conducted in developed countries with established cervical screening programs. The HPV FOCAL trial was rated as moderate quality. Applicability is discussed further in the following sections in relation to the specific questions addressed in this review.

The characteristics of the included trials is summarised in Table 81.

Table 81 Characteristics and quality of RCTs comparing HPV-based screening strategies with cytology (adapted and updated from Vesco et al. 2011 Tables 5a, 5b and 5c)

	NTCC Phase II	Finnish trial	NTCC Phase I	POBASCAM	Swedescreen	ARTISTIC	HPV FOCAL
Country	Italy	Finland	Italy	The Netherlands	Sweden	UK	Canada
New studies since AHRQ	Rossi et al. 2012 – no new data to contribute	Malila et al. 2012 – no new data to contribute Leinonen et al. 2012 – used to update AHRQ data	None	Rijkaart et al. 2012 – used to update AHRQ data	None	None	Oglivie et al. 2012 contributed all data
Screening approach – Round 1	HPV Primary: HPV (HC2) vs CC	LBC triage: HPV (HC2) with CC triage vs CC	Co-testing: HPV (HC2) + LBC vs CC	Co-testing: HPV (PCR) + CC vs CC	Co-testing: HPV (PCR) + CC vs CC	Co-testing: HPV (HC2) + LBC vs LBC	LBC triage: HPV (HC2) with LBC triage vs LBC with HPV (HC2) triage
Screening approach – Round 2	CC vs CC	NA	CC vs CC	HPV (PCR) +CC vs HPV (PCR) +CC	HPV (PCR) + CC vs CC	HPV (HC2) + LBC vs LBC	HPV (HC2) + LBC vs HPV (HC2) + LBC
Total randomised and screened	49,196	132,194	45,174	44,938	12,527	24,510	18,648
Ages recruited	25–60	25–65	25–60	30–56	32–38	20–64	25–65
Follow-up (yr)	3.5yr maximum after invitation to R2	3.6yr average (5 years maximum)	3.5yr maximum after invitation to R2	9yr	4.1yr mean	7 maximum	at least 12 months
Number of rounds	2	1	2	2	2	2[a]	2 (only 1 reported)
Round interval (y)	3	2–4	3	5	3	3	4
Threshold for colposcopy referral	IG: HPV+ CG: ASCUS+ (7 sites) or LSIL+ (2 sites)	IG: HPV+ & LSIL+ CG: LSIL+	IG: HPV+ or ASC-US+ (varied by age) CG: ASC-US+ (7 sites) or LSIL+ (2 sites)	IG: HSIL+ at any time; ASC-US+ at baseline and ASC-US+/HPV+ at 6 months; HPV+ at 6 months; HPV+ at second repeat smear at 18 months CG: HSIL+ at any time, ASC-US+ at baseline and 6 or 18 months	IG: test positive after repeat test at 12 months or positive by cytology as above CG: ASC-US+ referred to colposcopy in Stockholm, in other cities HSIL+	HSIL+	≥LSIL, ASC-US/HPV positive or persistent HPV positive all referred to colposcopy
Further testing protocol	HPV: none CC: baseline ASCUS (2 sites) →repeat cytology (timing NR) →if LSIL+ →colposcopy	IG: Baseline HPV+ and CC- or ASC-US → repeat screening (including HPV) at 12 months → repeat ASC-US, or three	IG: Baseline HPV+ and cytology negative → repeat HPV and cytology in one year, if HPV+ or ASC-US+ →	IG: normal cytology and HPV– recalled at 5 years. HPV– and LSIL or better recalled at 5 years. Repeat testing at 6 and 18 months for	IG: Baseline HPV+ (CC-) → repeat HPV and CC annually, persistent HPV+ (type specific) → colposcopy;	IG: HPV+ (LBC-): Baseline, repeat only HPV at 12 months, if HPV+ → patient choice of colposcopy at 12 months or repeat	CG: Repeat LBC screen at 2 years

	NTCC Phase II	Finnish trial	NTCC Phase I	POBASCAM	Swedescreen	ARTISTIC	HPV FOCAL
	Varied by age.	consecutive HPV+ results if CC- → colposcopy CG: Baseline ASC-US → repeat screening at 12 months (referral criteria NR)	colposcopy CG: Baseline ASC-US (2 sites)→ repeat cytology (timing NR), if LSIL+ → colposcopy	normal cytology/HPV+ and ASC-US+ CG: Normal cytology recalled at 5 years. ASC-US+ at baseline recalled at 6 and 18 months. Normal cytology at 6 and 18 months recalled at 18 months.	CG: Baseline ASC-US or LSIL → repeat CC only (timing and referral criteria NR)	HPV at 24 months, if HPV+ at 24 months → colposcopy; Same as below for LBC results CG: Baseline ASC-US or LSIL → repeat LBC at 6 and 12 months, if 3 consecutive ASC-US → colposcopy, or if 2 consecutive LSIL → colposcopy	
AHRQ quality rating	Fair	Fair	Fair	Fair	Fair	Fair	-

a. Extended follow-up including a 3rd round has been reported; however, the trial was designed to have 2 rounds.

Incidence and detection rates of CIN3+ and CIN2+

The relative detection ratios for CIN2+ and CIN3+ in the included trials are presented in Table 82 and Table 83, stratified by age for women aged greater or less than 30 or 35 years.

The rationale for cervical cancer screening is to identify and treat precancerous cervical lesions to prevent their progression to invasive cancer. Effectiveness depends, in part, on progression and regression rates of these lesions and their sojourn time (period from detectability until development into clinically manifest cancer), although these are not usually observable due to treatment. The effectiveness of a screening strategy can be demonstrated by reduced incidence of cervical cancer; however, conducting studies powered to detect this endpoint is generally not feasible. Therefore, CIN3+ (including the carcinoma in situ, adenocarcinoma or invasive cervical cancer) is considered an acceptable surrogate outcome. Reduced incidence of CIN3+ (which would be reflected in a reduced cumulative detection rate), may indicate that increased detection in early rounds is due to the detection and treatment of relevant progressive lesions (Arbyn et al. 2009).

Cumulative detection ratios for CIN3+ did not differ significantly between arms in four of five trials in the older age group; NTCC Phase II had a significantly increased cumulative detection of CIN3+.

In the first screening round three of seven trials demonstrated a significantly greater detection of CIN3+ in women aged ≥30–35 years with a HPV screening strategy compared with a cytology-only screening strategy. This was also observed in two of four trials reporting data for women aged ≤30–35 years. More trials reported increased detection of CIN2+ in the first screening round; five of seven trials in women ≥30–35 years and three of four in women aged ≤30–35 years. The details of these findings according to the different strategies and comparators are considered and discussed in detail in the corresponding sections below, relating to the specific review questions.

For five trials reporting data for the older age group, results of a second screening round were published and despite differences in referral thresholds and follow-up procedures, a consistent reduction in the relative detection ratio of CIN3+ was seen across all trials and was significant in three of four. A similar trend was observed for the detection of CIN2+, with a significant reduction in three of five trials. Increased detection of precancerous lesions in the first round followed by decreased detection in the second round may signal benefit due to increased detection and treatment of these lesions; however, the impact of the treatment of non-progressive lesions needs to also be considered.

These findings must be considered in light of the trade-off of increasing detection rates with the potential harms of screening. Measures to assess the potential harms of these screening strategies are presented in Table 84 and Table 85. Harms were generally poorly reported.

The colposcopy referral rates for the HPV-based strategies were greater in the intervention arm than the comparator arm in five of six studies in women aged ≥30–35 years, and in all five studies in women aged ≤30–35 years. In the younger women, the rates were elevated in both arms but particularly so for the HPV arm. Thus, while a HPV-based screening strategy tended to increase referrals to colposcopy, this effect was much greater in younger women (as indicated by the higher relative false positive proportions in this age group). Whether colposcopy referrals with a HPV-based strategy would decrease in latter rounds is not known.

Although these trials are not powered to detect a difference in cervical cancer incidence, there is a broad trend of a reduction in the number of cervical cancer cases detected in the second round in

the intervention arm (Table 81). This was significant in one trial (POBASCAM) (Rijkaart et al. 2012).

A modelled analysis of cervical cancer screening, diagnosis and treatment under IARC recommendations and considering the impact of HPV vaccination is necessary to explore the potential long-term benefits and trade-offs of these screening strategies in the Australian setting.

The results of these trials are discussed in further detail in the following sections relating to the specific questions addressed in this review.

Table 82 Relative detection ratios by screening round in HPV trials, women aged ≥30 or 35 years (adapted and updated from Vesco et al. 2011, Table 16a)

-	-	NTCC Phase II	Finnish trial	HPV Focal	NTCC Phase I	Swedescreen	POBASCAM**	ARTISTIC
		HPV primary	Reflex cytology	Reflex cytology	Co-testing	Co-testing	Co-testing	Co-testing
CIN2+	Round 1	↑ 2.13 (1.51–3.00)	↑ *1.60 (1.34–1.90)*	*1.46*[§]	↑ 1.78 (1.30–2.44)	↑ 1.51 (1.13–2.01)	↑ *1.25 (1.04–1.50)*	1.21 (0.91–1.60)
CIN2+	Round 2	↓ 0.25 (0.10–0.68)	*NA*	*NA*	0.59 (0.28–1.24)	↓ 0.58 (0.36–0.95)	*0.87 (0.70–1.07)*	↓0.63 (0.42–0.96)*
CIN2+	Cumulative	↑ 1.58 (1.16–2.13)	*NA*	*NA*	↑ 1.50 (1.13–1.98)	1.17 (0.92–1.49)	*1.08 (0.94–1.23)*	0.99 (0.83–1.19)*
CIN3+	Round 1	↑ 2.37 (1.44–3.89)	↑ ***1.56 (1.19**–2.05)*	*1.6*[§]	↑ 1.57 (1.02–2.43)	1.31 (0.93–1.86)	*1.15 (0.92–1.43)*	1.02 (0.71–1.47)
CIN3+	Round 2	↓ 0.23 (0.07–0.82)	*NA*	*NA*	0.46 (0.16–1.33)	↓ 0.53 (0.29–0.98)	↓ *0.72 (0.55–0.95)*	↓**0.53 (0.30**–0.96)*
CIN3+	Cumulative	↑ 1.57 (1.03–2.40)	*NA*	*NA*	1.30 (0.87–1.91)	1.04 (0.77–1.39)	*0.96 (0.81–1.13)*	0.85 (0.67–1.08)*
Invasive cervical cancer (Intervention:control)	Round 1	4:2	*NA*	*NR*	2:6	-	*12:6*	5:4
Invasive cervical cancer (Intervention:control)	Round 2	0:3	*NA*	*NA*	0:4	-	↓ *4:14*	3:0
Invasive cervical cancer (Intervention:control)	Cumulative	4:5 (ICC-AD)	*16:7 (ICC)*[†]	*NA*	2:10 (ICC-AD)	1:5 (ICC) 4:4 (ACIS-AD)	*16:20 (ICC)*	8:4*^ (ICC-AD)

* ARTISTIC CIN3+ and CIN2+ and ICC cases pooled across all ages at round 2; majority of participants (79%) were women >30 years

^ Note randomisation was 3:1; number of cancers reported for rounds 1 and 2 only

§ Not age stratified, includes women age 25–65, mean age 46, 93% of participtants were women >30 years. Data reported only as rates/1,000, contingency data not reported therefore confidence intervals not calculable (CIN2+: HPV arm 16.1 (13.2–18.9) vs control arm 11.0 (8.3– 13.7): CIN3+: HPV arm 8.0 (5.9–10.0) vs control arm 5.0 (3.1–6.8).

† These data are for women screened. For women who did not screen, the numbers were 8:18

** Entire trial population

Data in italics are updated from those reported in the AHRQ; data in bold are statistically significant. Arrows show the direction of effect.

Abbreviations: NA = not applicable, NR = not reported, ICC = invasive cervical cancer, AD = adenocarcinoma, ACIS = adenocarcinoma in situ.

Table 83 Relative detection ratios by screening round in HPV trials, women aged <30 or 35 years (updated and adapted from Vesco et al. 2011, Table 16b)

		NTCC Phase II	Finnish trial	NTCC Phase I	POBASCAM*	ARTISTIC
		HPV primary	Reflex cytology	Co-testing	Co-testing	Co-testing
CIN2+	Round 1	↑ 4.54 (2.95–6.99)	↑ *1.81 (1.45–2.26)*	↑ 1.99 (1.35–2.92)	*NR*	1.07 (0.83–1.38)
CIN2+	Round 2	↓ 0.40 (0.17–0.95)	*NA*	0.73 (0.34–1.60)	*NR*	NR
CIN2+	Cumulative	↑ 2.80 (1.98–3.95)	*NA*	↑ 1.63 (1.16–2.28)	*1.01 (0.81–1.26)*	NR
CIN3+	Round 1	↑ 4.00 (2.07–7.73)	↑ ***1.83 (1.21**–2.78)*	0.89 (0.51–1.57)	*NR*	0.92 (0.65–1.31)
CIN3+	Round 2	↓ 0.20 (0.05–0.93)	*NA*	1.00 (0.38–2.67)	*NR*	NR
CIN3+	Cumulative	↑ 2.19 (1.31–3.66)	*NA*	0.91 (0.56–1.48)	*0.97 (0.74–1.27)*	NR
Invasive cervical cancer (Intervention:control)	Round 1	1:0	*NA*	0:1	*NR*	NR
Invasive cervical cancer (Intervention:control)	Round 2	0:0	*NA*	0:2	*NR*	NR
Invasive cervical cancer (Intervention:control)	Cumulative	1:0 (ICC-AD)	*1:2 (ICC)*	0:3 (ICC-AD)	*NR*	NR

* Sub-group of trial population

Data in italics are updated from those reported in the AHRQ; data in bold are statistically significant. Arrows show the direction of effect.

Abbreviations: NA = not applicable, NR = not reported, ICC = invasive cervical cancer, AD = adenocarcinoma.

Table 84 Colposcopy referrals (%), relative false positive proportion and relative positive predictive value in HPV trials, women ≥30 or 35 years

-	-	NTCC Phase II	Finnish trial	HPV Focal	NTCC Phase I	Swedescreen	POBASCAM	ARTISTIC
-	-	HPV alone vs. Cyt	HPV reflex vs. Cyt	HPV reflex vs. Cyt	HPV co-test vs. Cyt	HPV co-test vs. Cyt	HPV co-test vs. Cyt	HPV co-test vs. Cyt
Colposcopy referrals (%)	Round 1	5.8 vs. 2.5	*0.92 vs. 0.99* *5.0* vs. 5.6**	*5.72§ vs. 3.32§*	10.6 vs. 3.0	NR	2.3† vs. 1.3† *5.0* vs. 2.9**	4.9 vs. 4.1
	Round 2	NR	*NA*	*NA*	NR	NR	1.3† vs.1.9†	160 (N) vs. 42 (N)
	Cumulative	NR	*NA*	*NA*	NR	NR	3.4† vs. 2.8†	6 vs. 4.9
Relative false positive proportion for CIN3+	Round 1	2.42 (2.15– 2.72)	*0.88 (0.84–0.92)*	*NR*	NR	NR	NR	1.67 (1.52–1.84)
	Round 2	NR	*NA*	*NA*	NR	NR	NR	2.11 (1.82–2.45)^
	Cumulative	NR	*NA*	*NA*	NR	NR	NR	1.86 (1.74–1.98)^
Relative PPV for CIN3+	Round 1	0.86 (0.49–1.52)	*1.74 (1.33–2.29)*	*0.93*	0.34 (0.21–0.54)	NR	NR	0.63 (0.44–0.90)
	Round 2	NR	*NA*	*NA*	NR	NR	NR	0.32 (0.18–0.55)^
	Cumulative	NR	*NA*	*NA*	NR	NR	NR	0.54 (0.44–0.66)^

* Referred for follow-up testing

^ All age data reported; majority of participants (79%) were women >30 years

§ Not age stratified, includes women age 25–65, mean age 46, 93% of participants were >30 years

† Data from 18,403 with sufficient follow-up from 2007 analysis (Bulkmans et al. 2007)

Data in italics are updated from those reported in the AHRQ Arrows show the direction of effect.

Abbreviations: NA = not applicable, NR = not reported, cyt = cytology, PPV = Positive predictive value, HPV = human papillomavirus.

Table 85 Colposcopy referrals (%), relative false positive proportion and relative positive predictive value in HPV trials, women <30 or 35 years

		NTCC Phase II	Finnish trial	NTCC Phase I	POBASCAM	ARTISTIC
		HPV alone vs. Cyt	HPV reflex vs. Cyt	HPV co-test vs. Cyt	HPV co-test vs. Cyt	HPV co-test vs. Cyt
Colposcopy referrals (%)	Round 1	13.1 vs. 3.6	*2.6 vs. 1.9* *15.1* vs. 6.0**	11.9 vs. 4.1	*11.3* vs. 5.5**	13.9 vs. 9.6
	Round 2	NR	*NA*	NR	-	-
	Cumulative	NR	*NA*	NR	-	17.1 vs.12.0
Relative false positive proportion for CIN3+	Round 1	4.02 (3.46–4.67)	*2.24 (2.07–2.41)*	NR	NR	2.02 (1.79–2.28)
	Round 2	NR	*NA*	NR	NR	-
	Cumulative	NR	*NA*	NR	NR	-
Relative PPV for CIN3+	Round 1	0.66 (0.31–1.40)	*0.83 (0.55–1.24)*	0.80 (0.55–1.18)	NR	0.50 (0.36–0.69)
	Round 2	NR	*NA*	NR	NR	-
	Cumulative	NR	*NA*	NR	NR	-

* Referred for follow-up testing

Data in italics are updated from those reported in the AHRQ

Arrows show the direction of effect.

Abbreviations: NA = not applicable, NR = not reported, cyt = cytology, PPV = Positive predictive value, HPV = human papillomavirus.

HPV testing for primary screening

Primary research question 3

What is the comparative safety, effectiveness and cost-effectiveness of HPV testing as the primary screening test in women aged 25 to 65 years every five years, compared with the protocol used in the current Australian cervical screening program?

Summary 10: Summary of effectiveness—HPV primary screening (without triage or co-testing)

Longitudinal outcomes from RCTs

In one RCT (NTCC Phase II), HPV primary screening (without cytology co-testing or triage) led to an increase in the detection of CIN2+ and CIN3+ in comparison to conventional cytology screening, which was greater in women <35 years than women ≥35 years. This trial was the only HPV trial that had a significant increase in the cumulative detection of CIN3+, which was observed in both younger and older age groups. It also had the greatest increase in cumulative detection of CIN2+. The extent to which this is due to the increased detection of clinically significant disease is unclear. However, the increase in the detection ratios in unvaccinated women younger than age 35 in comparison to older women suggests that an excess of regressive lesions was identified in the HPV screening arm in this age group.

At round one, there were significantly more colposcopies in the HPV arm of the trial and the false positive proportion was significantly greater for the HPV strategy. The relative increase in CIN2+ and CIN3+ detection, the relative increase in colposcopy referral, and the relative false positive proportion was greater in unvaccinated younger women (<35 years) suggesting that HPV testing detects more clinically insignificant lesions in this age group.

This study was conducted in unvaccinated women and no triage strategy was employed. It is expected that the number of HPV positive tests and the number of colposcopies performed would be less in vaccinated women. A modelled analysis of cervical cancer screening, diagnosis and treatment using the IARC recommendations in HPV vaccinated and unvaccinated populations has been conducted to explore the long-term benefits and potential harms of these alternative screening strategies in the Australian setting.

Cross-sectional accuracy

A high-quality meta-analysis indicated that a primary HPV screening strategy is much more sensitive than cytology at the thresholds of LSIL+ and ASC-US+ for the detection of CIN2+ and CIN3+ with a small loss in specificity. Two studies included in the AHRQ review reporting on absolute accuracy measures also indicated that HPV screening was more sensitive than screening with cytology, but less specific. Relative specificity is reduced to a greater degree in younger women in one study identified by the AHRQ review.

Conclusions

- HPV testing alone (without triaging) for primary screening is more sensitive and less specific than cytology screening for CIN2+ and CIN3+.
- HPV testing alone (without triaging) for primary screening increases colposcopy referrals in comparison to cytology-based screening.

- The high rate of detection of CIN2+ and CIN3+ in unvaccinated women younger than age 35 suggests that an excess of regressive lesions is identified with HPV screening in this age group (NTCC HPV testing alone trial) in unvaccinated populations.

MSAC has previously assessed the safety, effectiveness, and cost-effectiveness of HPV testing as a stand-alone screening test (or combined with screening by cytology) in 2003 (MSAC Reference 12d, see Appendix J). In 2003, MSAC concluded that there was insufficient evidence to support public funding of HPV testing either as a stand-alone screening test or as an adjunct to conventional cytology.

The current review updates the AHRQ review (Vesco et al. 2011). No new studies were identified. The AHRQ included one trial, NTCC Phase II, in a developed country which assessed HPV primary screening using HC2 compared to conventional cytology. The threshold for referral to colposcopy in this study varied according to site for the cytology arm (ASCUS+ at seven sites; LSIL+ at two sites), which is lower than that in Australia (HSIL+). All women in the intervention arm with a HPV positive screening test were referred to colposcopy. The trial was conducted over two screening rounds; in the second round screening was conventional cytology for all women. The screening interval was three years which differs to both the current Australian recommendation of two years for conventional cytology and the proposed screening interval of five years for HPV primary screening.

The AHRQ rated the trial as fair quality. Limitations of the study include lack of blinding of participants, colposcopists and histologists, lack of biopsies for negative colposcopies and that non-compliant women in round one were not invited back for round two (2.8% HPV vs 0.7% CC). The trial is of high applicability to the current review.

An extract from the AHRQ report (Vesco et al. 2011) executive summary is provided below; these comments pertain to women older than 35 years of age, after two screening rounds in the NTCC Phase II trial.

NTCC Phase II, women >35 years (Vesco et al. 2011):

'After two rounds of screening in NTCC Phase II (one round of HPV screening) and a median of 3.5 years of follow up, cumulative detection of CIN3+ (CIN3, AIS, or ICC) was increased in 17,724 women screened with HC2 relative to 17,747 women screened with cytology alone (55 vs. 35 CIN3+ lesions; RR, 1.57 [95% CI, 1.03 to 2.40]), with about the same number of invasive cancer cases detected in both arms (HC2 arm: 4 ICC/AIS cases; cytology alone: 5 ICC/AIS cases).'

A similar pattern was observed among women younger than 35 years. An extract from the AHRQ report executive summary is provided below.

NTCC Phase II, women <35 years (Vesco et al. 2011):

"The pattern of results in 13,725 younger women was similar to older women, but with a much higher rate of colposcopy referrals after HC2 screening. After two rounds of screening in NTCC Phase II, cumulative detection of CIN3+ also increased in younger women screened with HC2 relative to cytology alone (47 vs. 21 CIN3+ lesions; RR, 2.19 [95% CI, 1.31 to 3.66]), with few ICC cases detected in either arm (HC2 arm: 1 case; cytology arm: 0 cases)."

The significant decrease in the detection of both CIN2+ (≥35: RR 0.25 [95% CI 0.10 – 0.68]; >35: RR 0.40 [95% CI 0.17 – 0.95]) and CIN3+ (≥35: RR 0.23 [95% CI 0.07 – 0.82]; >35: RR 0.20 [95% CI 0.05 – 0.93]) (Table 82 and Table 83) in the second round for all women is indicative of a preventative benefit from HPV screening in the first round.

However, the NTCC Phase II trial was the only HPV trial that demonstrated a significant increase in cumulative CIN3+. This was observed in both younger and older age groups. As the second screening round in this trial was cytology in both arms and the screening interval was relatively short (3 years), it is not possible to determine to what extent the excess detection of CIN3+ in this trial is due to earlier detection of lesions (gains in lead time) or excess detection of regressive lesions ('overdiagnosis'). However, the relative detection ratios of both CIN2+ (RR 4.54 [95% CI 2.95–6.99]) and CIN3+ (RR 4.00 [95% CI 2.07–7.73]) in younger women were elevated in the first round (Table 83) and compared to that in older women (CIN2+: RR 2.13 (95% CI 1.51–3.00); CIN3+: RR 2.37 (95% CI 1.44–3.89)) (Table 82). This increase in the detection ratios in women younger than age 35 in comparison to older women suggests that an excess of regressive lesions was identified in the HPV screening arm in this age group.

The trade-offs associated with screening in the NTCC trial are summarised in Table 84 and Table 85. The NTCC trial did not report cumulative results for PPV, false-positive rates or colposcopy; therefore, it is difficult to assess the harms of primary HPV screening. For women aged 35 and over, baseline colposcopy referral rates were higher in the HC2 arm than the cytology arm (HC2 5.8% versus CC 2.5%; RR 2.37 (2.13–2.65), P<0.001) but are not reported for the entire first round or second round of screening. However, due to the low threshold for colposcopy referral (lower than that in Australia) this is likely to approximate the entire first round (Vesco et al. 2011). For women under aged 35, baseline colposcopy was much higher for HPV-screened women compared to those screened with cytology (HC2 13.1% vs CC 3.6%; RR 3.29 (2.88–3.75), P<0.001) and was notably higher than that in older women. Similarly, the false positive proportion was significantly greater in the HPV arm than the CC arm in round one (>35 years: RR 2.42 (95% CI 2.15–2.72)), and this was significantly greater in younger than older women (>35 years: 4.02 (95% CI 3.46–4.67). Data were not reported for round two.

A modelled analysis of cervical cancer screening, diagnosis and treatment under IARC recommendations and considering the impact of HPV vaccination is necessary to explore the potential long-term benefits and trade-offs of primary HPV screening in the Australian setting.

As RCT level evidence was identified to address this research question, primary studies providing a lower level of evidence were excluded. However, a summary of data from the most recent systematic review reporting cross-sectional accuracy (Arbyn et al. 2012) is included in addition to data reported on absolute accuracy in the AHRQ HTA (Vesco et al 2011).

Meta-analyses

Arbyn conducted a meta-analysis of the relative accuracy of HC2 as a primary screening test to detect CIN2+ and CIN3+. The meta-analysis included 49 studies, the majority of which used the HC2 or GP5+/6+ PCR test; only results for HC2 are discussed here. The relative sensitivity and specificity at the cytology thresholds of LSIL+ and ASC-US+ for the detection of CIN3+ is shown in Table 86. Studies with comparators of LBC or conventional cytology were pooled. For the detection of CIN3+, the relative sensitivity of 1.36 (95% CI; 1.21–1.53, range 0.97–2.32) (13 studies) was significant showing higher sensitivity for HC2 than cytology and the relative specificity was 0.93 (95% CI; 0.91–0.96, range 0.84–1.03) (12 studies). The pooled relative sensitivity of HC2 (at the standard threshold of ≥1RLU) compared to cytology at the threshold of LSIL+ for the detection of CIN2+ was 1.40 (95% CI 1.27–1.54, range 1.09–2.37) (20 studies) while the relative specificity was 0.92 (95% CI 0.90–0.94, range 0.67–1.03) (19 studies). These data are in accord with those reported in round one of the NTCC Phase II trial. No age-specific data were reported.

Cross-sectional / absolute test accuracy

The AHRQ HTA included six community-based studies in developed countries which reported on the absolute test performance of HPV testing alone compared with cytology. Only two of these studies reported sensitivity and specificity for the detection of CIN3+ and the AHRQ only reported data for the cytology threshold of ASC-US+ which is a lower threshold for repeat testing or colposcopy than is used in Australia. Nevertheless, these two studies were in broad agreement with the findings of the Arbyn et al. meta-analysis (2012); markedly higher sensitivity for HC2 with reduced specificity for CIN3+. Specificity of HC2 relative to cytology was decreased to a much greater degree in young women in one of these studies. The cross-sectional accuracy of these two studies for the detection of CIN3+ is shown in Table 87.

Table 86 Relative accuracy of HC2 screening strategies for the detection of CIN3+ and CIN2+ as reported in Arbyn et al. 2012

Intervention	Comparator	Cytology threshold	Outcome	N of studies (Sn/Sp)	Relative sensitivity (95% CI)	Range	Relative specificity (95% CI)	Range
HC2	Cytology	ASC-US+	CIN3+	20/18	1.27 (1.12–1.44)	0.97–2.63	0.97 (0.96–0.99)	0.88–1.10
HC2	Cytology	ASC-US+	CIN3+	8/6*	1.43 (1.15–1.77)	1.01–2.12	0.97 (0.96–0.98)	0.93–1.00
HC2	Cytology	LSIL+	CIN3+	13/12	1.36 (1.21–1.53)	0.97–2.32	0.93 (0.91–0.96)	0.84–1.03
HC2 & cytology co-test	Cytology	ASC-US+	CIN3+	10/9	1.33 (1.29–1.37)	1.02–2.18	0.92 (0.91–0.93)	0.85–0.96
HC2 & cytology co-test	HC2	ASC-US+	CIN3+	6/6^	1.02 (1.01–1.03)	1.04–1.04	0.93 (0.92–0.95)	0.81–0.99
HC2	Cytology	ASC-US+	CIN2+	28/25	1.23 (1.15–1.31)	0.91–2.93	0.97 (0.96–0.98)	0.86–1.10
HC2	Cytology	ASC-US+	CIN2+	12/10*	1.37 (1.22–1.54)	1.06–2.25	0.97 (0.96–0.98)	0.93–1.00
HC2	Cytology	LSIL+	CIN2+	2,019	1.40 (1.27–1.54)	1.09–2.37	0.92 (0.90–0.94)	0.67–1.10
HC2 & cytology co-test	Cytology	ASC-US+	CIN2+	13/13	1.42 (1.36–1.48)	1.06–2.30	0.94 (0.93–0.94)	0.89–0.96
HC2 & cytology co-test	HC2	ASC-US+	CIN2+	10/10^	1.05 (1.04–1.07)	1.00–1.19	0.95 (0.94–0.96)	0.81–0.99

* Restricted to studies in North America or Europe. Also excludes Kitchener et al. 2009; basis unclear
^ Excludes Sankaranarayanan et al. 2005, basis unclear
Abbreviations: HC2 = hypbrid capture 2, Sn = Sensitivity, Sp = Specificity, CI = confidence interval

Table 87 Cross-sectional accuracy for the detection of CIN3+ as reported in Vesco et al. 2011

Study	Age group	Outcome	Sample size	Sensitivity HC2	Sensitivity Cytology (ASC-US+)	Sensitivity HC2 and cytology	Specificity HC2	Specificity Cytology (ASC-US+)	Specificity HC2 and cytology
Kulasingam 2002	**Women ≥30 years**	CIN3+	774	86.0 (59.7– 96.9)	49.7 (32.9–71.5)	49.7 (32.9–71.5)^	83.0 (76.8– 87.1)	86.4 (84.8–88.1)	94.7 (92.8–96.1)
Kulasingam 2002	Women <30 years	CIN3+	3,301	92.5 (83.5– 97.3)	65.4 (51.9–79.1)	64.0 (51.1–77.6)^	70.1 (66.5 – 73.1)	81.5 (80.7–82.3)	87.6 (86.7–88.4)
Petry 2003	**Women ≥30 years**	CIN3+	7,908	97.3 (83.2–99.6)	46.0 (30.8–61.9)	100 (93.7–100)*	95.2 (93.4–96.5)	98.0 (96.7–98.8)	94.9 (93.1–96.2)
Kulasingam 2002	**Women ≥30 years**	CIN2+	774	62.7 (31.4–93.2)	38.3 (19.3–63.3)	38.3 (19.3–63.3)^	83.0 (76.6–87.2)	86.4 (84.7–88.3)	95.0 (93.0–96.4)
Kulasingam 2002	Women <30 years	CIN2+	3,301	73.5 (53.3–87.7)	50.1 (35.2–62.2)	47.9 (34.1–60.0)^	71.1 (67.3–74.0)	82.1 (81.3–83.0)	88.3 (87.4–89.2)
Petry 2003	**Women ≥30 years**	CIN2+	7,908	97.8 (86.3–99.7)	43.5 (30.0–58.0)	100 (93.7–100)*	95.3 (93.5–96.6)	98.0 (96.7–98.8)	93.8 (91.8–95.3)

* Positive test defined as either positive HC2 or positive cytology (PapIII+)
^ Positive test defined as both HC2 positive and cytology ASC-US+

LBC and HPV co-testing for primary screening

Secondary research question 3.1

What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a co-test in the above scenario, compared with the primary HPV testing?

Summary 11: Summary of effectiveness – HPV and cytology co-testing

Longitudinal outcomes from RCTs

Co-testing involves screening every woman with both a HPV test and a cytology test. The definition of a positive test (ie as both tests positive or either test positive) and the screening pathway following these tests varied greatly across the included studies.

Four RCTs were included, three of which provided a comparison to conventional cytology (NTCC Phase I, Swedescreen and POBASCAM) and one to LBC (ARTISTIC). None of the trials showed any difference in the cumulative detection of CIN3+. One RCT (NTCC Phase I) comparing HPV and LBC co-testing to conventional cytology had increased cumulative detection of CIN2+ and suggests that excess regressive disease is being detected in this trial in both younger and older age groups. This trial had a co-testing strategy which had a lower threshold for referral to colposcopy than the other trials and significantly more women were referred to colposcopy in the co-testing arm than the cytology arm.

Across trials there tended to be an increase in the detection of CIN2+ in round one followed by a decrease in detection of CIN3+ in round two.

A comparison of the intervention arms of NTCC Phase I (co-testing) and II (HPV primary) trials indicates similar sensitivity of these strategies, with higher colposcopy rates in the co-testing strategy. Such indirect comparisons are prone to bias but suggest no clear advantage of co-testing over HPV testing alone.

Cross-sectional accuracy

A high-quality meta-analysis indicated that co-testing is more sensitive, but less specific, than cytology alone at the threshold of pLSIL+ for the detection of CIN3+ when abnormalities on either test are considered positive. However, this gain in sensitivity is similar in magnitude to that for HPV testing alone.

In one study included in the AHRQ HTA, a strategy in which both tests are required to show an abnormality for a positive test result demonstrated similar sensitivity to cytology alone (but lower than that for HPV testing alone) and increased specificity compared to either test alone. This approach is considered equivalent to a triage strategy.

Conclusions

- Co-testing (adjunctive or dual HPV and cytology testing) for either HPV or cytology positivity is marginally (but significantly) more sensitive than HPV testing alone.
- Co-testing for either HPV or cytology positivity is significantly less specific than HPV testing alone.

- HPV and cytology co-testing does not demonstrate a clear advantage over HPV testing alone (based on indirect comparisons which are prone to bias).

MSAC last considered evidence for HPV and cytology co-testing in May 2003. At this time only one study (Kulasingam et al. 2002) met the inclusion criteria for review. This study provided data on the effectiveness of HPV and LBC co-testing in comparison with LBC alone. MSAC concluded that there was insufficient evidence that HPV co-testing added advantages to screening by cytology alone at that time.

The current review updates the AHRQ review (Vesco et al. 2011). No new studies were identified; however, an update to the POBASCAM trial (Rijkaart et al. 2012) was identified reporting full follow-up for round two. Overall, the addition of the final results of the POBASCAM trial do not significantly change the body of data presented in the AHRQ report. The AHRQ review included four RCTs of HPV and cytology co-testing; NTCC Phase I, POBASCAM, Swedescreen and ARTISTIC, with marked differences in protocols across the trials. All trials were conducted in developed countries with established screening programs.

The experimental and control arms of the trials were HC2 and LBC versus conventional cytology (NTCC Phase I), HC2 plus LBC versus LBC (ARTISTIC) and PCR (GP5+/6+) plus CC versus CC (POBASCAM and Swedescreen.) Although this review considers conventional and liquid-based cytology as equivalent for the purposes of evaluating the benefits of HPV screening, the specific question addressed in this review is HPV testing (HC2) plus LBC versus conventional cytology; NTCC Phase I specifically addressed this question. Round 2 procedures for each trial also varied, for Swedescreen and ARTISTIC it was the same as round one, for NTCC Phase I it was conventional cytology for both arms and for POBASCAM it was PCR (GP5+/6+) plus CC for both arms.

In respect to follow-up testing procedures and referral to colposcopy, three of the four trials (POBASCAM, Swedescreen, ARTISTIC) had a "high threshold" for colposcopy, which was the same as that in Australia (ie HSIL+), although for Swedescreen this varied by site. For NTCC Phase I, the threshold was lower than Australian practice and varied by site; ASC-US+ (seven sites) or LSIL+ (two sites). The three trials with a high cytology threshold (POBASCAM, Swedescreen, ARTISTIC) also had a high threshold for the HPV testing arm; immediate referral to colposcopy was dependent on cytology alone with HPV positive results triggering repeat testing (6 or 12 months) and referral to colposcopy following persistent HPV positivity and/or abnormal cytology. The HPV co-testing arm of the NTCC Phase I trial followed a similar strategy in women younger than aged 35, but for older women a HPV positive test triggered immediate referral to colposcopy. Therefore, while the NTCC Phase I trial is most applicable to the research question posed in this review, it had a low referral threshold for colposcopy which does not reflect current Australian practice. In contrast, the other three trials had repeat testing procedures which are more likely to reflect Australian practice.

The AHRQ notes the duration and completeness of follow-up and lack of adherence to trial protocols are limitations across these four trials. For example, 29% of the sample in ARTISTIC had less than the minimal (2.5 year) follow-up after round two, and follow-up after round two in Swedescreen averaged less than one year. The updated results from POBASCAM included in this review improve the limitations for this trial. However, only 50–60% of participants complied with repeat testing procedures and 20% had opportunistic screening outside the trial. The AHRQ rated all four trials as fair quality and these issues as well as lack of blinding were the key quality issues.

In women more than 30 or 35 years of age, a co-testing strategy significantly increased the relative detection ratios of CIN2+ in round one in three of four trials (NTCC Phase I, Swedescreen and POBASCAM), and of CIN3+ in one trial (NTCC Phase I; Table 82). In all trials the detection of CIN2+ and CIN3+ was reduced in round two. This was significant for both outcomes in the two trials, Swedescreen and ARTISTIC, that maintained the same screening strategy in round one and round two (co-testing vs cytology (Table 82)). In the two trials in which screening strategies in round two were the same for both trial arms, only the detection of CIN3+ in the POBASCAM trial was statistically significantly less than unity.

For NTCC Phase I there was a significantly increased cumulative detection of CIN2+. Given that CIN3+ was not reduced cumulatively, this may indicate that regressive lesions are being detected in excess numbers in this trial. None of these co-testing trials demonstrated an increase in cumulative CIN3+ detection, which was observed in the NTCC Phase II trials of HPV testing alone. However, such indirect comparisons are prone to bias.

Only NTCC Phase I and ARTISTIC included women younger than age 30 or 35 years and complete age-specific data were only reported for NTCC Phase I. This trial found no change in CIN3+ detection rates in any round or cumulatively in younger women. However, CIN2+ detection was greatly increased in round one and cumulatively, suggesting that excess regressive disease is being detected in these women.

Cumulative invasive cancer detection tended to be higher in the cytology arms of co-testing trials; however, due to low numbers and the described issues with follow-up, these findings should be considered cautiously. The full follow-up from the POBASCAM trial which reported cancer detection for each round did show an increased detection of cancer in the experimental arm in round one followed by decreased detection in round two, and an overall cumulative reduction. The round two results were significant (4/19,579 vs 14/19,731; 0.29, 0.10–0.87; p=0.031).

The findings of this review including updated data are in agreement with those of the AHRQ report (Vesco et al. 2011). An extract from the AHRQ report executive summary is provided below. These comments relate to 82,390 European women aged 30–64 years from the four included RCTs.

AHRQ executive summary, any HPV co-testing, 4 trials (Vesco et al. 2011)

Women aged ≥30–35 years
"Cumulative relative CIN3+ detection was the same between HPV-cytology co-testing and cytology alone for women greater than 30-35 years of age after two screening rounds in all the RCTs, and most co-testing trials report differences in round-specific relative CIN detection (e.g., more CIN2+ with co-testing after Round 1, and less CIN3+ with co-testing after Round 2)."

Women aged <30–35 years
"No impact on CIN3+ in any round or cumulatively was seen in younger women. CIN2+ detection was relatively greater after Round 1 and cumulatively with co-testing perhaps reflecting overdiagnosis of regressive disease."

Referral to colposcopy and other harms were poorly reported across all four trials (Table 84 and Table 85). Cumulative colposcopies were higher in the intervention arm of ARTISTIC for both older (6.0% vs 4.9%; RR 1.21 (1.05–1.29), P<0.01) and younger women (17.1% vs 12.0%; RR 1.92 (1.59–2.31), P<0.001), with younger women having a higher rate overall. The false positive rates

were significantly higher in the co-testing arm than in the cytology arm in this trial across both rounds and cumulatively (round one false positive proportion 1.67 [95% CI 1.52–1.84] in women ≥30 years; 2.02 [95% CI 1.79–2.28] women <30 years). The increase in false positive rates was, however, less than that observed in the NTCC Phase II trial of HPV testing alone in both younger and older women (round one false positive proportion 2.42 [95% CI 2.15–2.72] in women ≥35 years; 4.02 [95% CI 3.46–4.67] women <35 years); however, such indirect comparisons must be considered with caution.

Colposcopy referral rates are highly dependent on the referral protocol used (ie both tests positive vs one test positive vs either test positive) in individual trials. Colposcopies were not reported in the final results of POBASCAM, but earlier interim analysis demonstrated relatively low rates and a small increase in the intervention arm compared to the comparator arm (round 1: co-testing 2.3% (2.0–2.7) vs cytology 1.3% (1.1–1.6), P<0.0001; round two: co-testing 1.3% (1.0–1.6) vs cytology 1.9% (1.6–2.2), p=0.003); this trial had a high referral threshold for colposcopy. In NTCC Phase I, with a low colposcopy threshold, immediate colposcopies were much higher with co-testing than with cytology alone (women ≥30–35 years co-testing 10.6% (10.2–11.1) vs conventional cytology 3.0% (2.7–3.3), P<0.0001; women <30–35 years co-testing 11.9% (11.1–12.7) vs conventional cytology 4.1% (3.6–4.6), P<0.0001). Round one colposcopy rates were higher in the intervention arm of NTCC Phase I than Phase II for women older than 35 (10.6% vs 5.8%) reflecting the protocols in which women with either test positive were referred to colposcopy (HPV+ or ASC-US+, NTCC Phase I) compared to only a HPV+ test when this was the only test used (NTCC Phase II).

The comparison of the intervention arms of NTCC Phase I and II is the best indirect measure of the added value of co-testing to HPV screening as the trials differed only in whether HPV was used as a stand-alone test or in combination with cytology. Cumulative detections rates for CIN2+ and CIN3+ were similar (CIN3+: 1.50 (1.13–1.98) Phase I vs 1.58 (1.16–2.13) Phase II; CIN2+: 1.30 (0.87–1.91) Phase I vs 1.57 (1.03–2.40) Phase II) in women aged ≥35, despite the greater colposcopy rate in the co-testing arm, suggesting that co-testing increases false positive rates without increasing sensitivity.

In summary, none of these four trials (including the one trial of HPV and LBC versus CC) provided evidence that co-testing significantly reduces the cumulative incidence of CIN3+ over progressive screening rounds, compared to cytology alone. However, all trials showed a reduced detection rate ratio of CIN3+ in round two which ARTISTIC and POBASCAM used as the primary outcome for power calculations.

As RCT level evidence was identified to address this research question, studies providing a lower level of evidence were excluded. However, a summary of data from the most recent systematic review reporting cross-sectional accuracy (Arbyn et al. 2012) is included in addition to data reported on absolute and relative accuracy in the AHRQ HTA.

Meta-analyses

In the high-quality meta-analysis by Arbyn et al. (2012), the pooled relative sensitivity of combined testing compared to cytology alone at a cytology threshold of ASC-US+ was significantly increased (for the detection CIN2+ was 1.42 (95% CI, 1.36–1.48) (13 studies) and for the detection of CIN3+ was 1.33 (95% CI, 1.29–1.37) (10 studies)). The pooled relative specificities were significantly reduced at 0.94 (95% CI, 0.93–0.94; CIN3+, 13 studies) and 0.92 (95% CI, 0.91–0.93; CIN2+, nine studies) respectively. Thus co-testing was more sensitive but less specific than cytology testing alone.

The relative sensitivity of combined testing compared to HPV testing (HC2) alone was 1.05 (95% CI, 1.04–1.07) (10 studies) for the detection of CIN2+ and 1.02 (95% CI, 1.01–1.03) (six studies) for the detection of CIN3+. The specificities were 0.95 (95% CI, 0.95–0.96) (10 studies) and 0.93 (95% CI, 0.92–0.95) (six studies) respectively. This indicates that co-testing provides a small, but significant increase in sensitivity compared to HPV testing alone, accompanied with a decrease in specificity. Arbyn did not include a meta-analysis for cytology thresholds other than ASC-US+ and did not discuss whether or not the studies included in the meta-analysis used differing criteria to define a positive co-test.

Cross-sectional / absolute test accuracy

The AHRQ report included four cohort studies which reported the absolute accuracy of combined HPV and cytology compared to cytology or HC2 alone (Table 87). The studies all defined a positive co-test result differently. Only two of these studies reported sensitivity and specificity for the detection of CIN3+ and the AHRQ only reported data for the cytology threshold of ASC-US+ which is a lower threshold for repeat testing or colposcopy than is used in Australia.

Either test positive

In one of these studies a positive co-test was defined as either test being abnormal. In this study the sensitivity of co-testing was significantly increased compared to cytology alone (100% (95% CI 93.7–100) vs 46% (95% CI 30.8–61.9) while the specificity was significantly decreased (94.9% (95% CI 93.1–96.2) vs 98.0 (95% CI 96.7–98.8). The accuracy was similar to that of HC2 testing alone compared to cytology, giving only a small non-significant increase in sensitivity for HPV co-testing in comparison to HC2 alone, with a small but significant reduction in specificity.

Both tests positive

In the second study (Kulasingam et al. 2002), both tests were required to be positive for a positive co-test. Results were similar in this study for women aged less than 30 years and women aged more than 30 years (see Table 87). The sensitivity of the co-test for the detection of CIN3+ was similar to that of cytology alone (women <30: 64.0% (95% CI, 51.1–77.6) vs 65.4% (95% CI, 51.9–79.1); women ≥30: 49.7 (95% CI, 32.9–71.5) vs 49.7 (95% CI , 32.9–71.5) while the specificity was increased (women <30 87.6% (95% CI, 86.7%–88.4%) vs 81.5% (95% CI, 80.7–82.3); women ≥30: 94.7 (95% CI, 92.8–96.1) vs 86.4 (95% CI, 84.8–88.1)). Compared to HC2 alone sensitivity was significantly reduced and specificity increased.

Therefore, the relative accuracy of a co-testing strategy depends on the strategy itself; if abnormality of either test is considered a positive test result then the co-test gives no significant increase in sensitivity over HC2 alone, but a small but statistically significant reduction in specificity. On the other hand, if both tests are required to be abnormal for a positive test then the sensitivity is not significantly different to that of cytology alone, but there is an increase in specificity. Such a strategy is similar to testing with either cytology or HC2 alone followed by triage of positive tests using the other test.

LBC triage of HPV testing for primary screening

Secondary research question 3.2

What is the comparative safety, effectiveness and cost-effectiveness of manually read LBC or automated image analysis LBC as a reflex test to triage women with positive HPV test results, compared with primary HPV testing?

Summary 12: Summary of effectiveness—cytology triage of HPV testing

Longitudinal outcomes from RCTs

No RCTs were identified which provided a comparison of HPV primary screening with and without LBC triage.

Two RCTs of HPV with LBC triage providing comparisons to conventional cytology alone or LBC with HPV triage were included. Both trials only reported on one screening round and were conducted in unvaccinated populations.

The Finnish trial demonstrated that HPV with LBC triage increases the detection of CIN2+ and CIN3+ lesions in comparison to CC.

- This was greater in women less than 35 years.
- Younger women also had a greater number of referrals to colposcopy at the index screen (at a referral threshold of LSIL+) and repeat testing, with higher rates in the HPV with LBC triage arm.
- In contrast, colposcopy referrals (at a referral threshold of LSIL+) and repeat testing rates were similar for both testing strategies in older women (≥35years).

In the HPV FOCAL trial, the detection of CIN2+ and CIN3+ lesions was non-significantly higher in the HPV with LBC triage arm compared to the LBC with HPV triage arm after the first round. Referral to colposcopy was increased in the HPV with LBC triage arm of the trial, but these data were not age-stratified.

These studies were conducted in unvaccinated women. It is expected that the number of positive tests and the number of colposcopies would decline in vaccinated women. A modelled analysis of cervical cancer screening, diagnosis and treatment using the IARC recommendations in HPV-vaccinated and unvaccinated populations has been conducted to explore the potential long-term benefits and harms of these alternative screening strategies in the Australian setting.

Cross-sectional accuracy

One study (Kulasingam et al. 2002) of HPV and cytology co-testing, in which an abnormal result on both tests is required for a positive result, provides accuracy data indicative of a triage approach. This study indicated that the sensitivity of the co-testing strategy for CIN3+ was lower than that of HPV testing alone while the specificity was increased. Results were similar for women aged less and more than 30 years.

Conclusions

- Cytology triage of HPV primary testing has a higher colposcopy and retesting rate than cytology screening alone (at a referral threshold of LSIL+) in unvaccinated women <35 years.
- Cytology triage of HPV primary testing has a similar colposcopy and retesting rate to cytology screening alone (at a referral threshold of LSIL+) in women ≥35 years.

MSAC has not previously considered evidence for LBC triage of HPV testing for primary screening.

The current review updates the AHRQ review (Vesco et al. 2011). Two studies of HPV with LBC triage with different comparators were included for review. The AHRQ HTA identified one RCT (the Finnish study) of HPV with LBC triage in comparison with conventional cytology. Two further publications (Malila 2012; Leinonen 2012) from this RCT were identified in the update search, one of which (Leinonen 2012) provided updated data for inclusion in this review, increasing the number of women in the analyses from 71,337 to 132,194. However, results from a planned second screening round (five-year screening interval) are not yet reported. A second RCT (HPV FOCAL) was also identified in the update screen. This trial reports on a comparison of HPV with LBC triage in comparison to LBC with HPV triage. This trial was not included in the AHRQ review as no results were available at that time; the new publication (Ogilvie et al. 2012) provides results from the first round of screening and is included in the current review.

No RCTs were identified that directly addressed the research question of LBC triage of HPV testing versus HPV testing alone. Co-testing strategies where only high-grade cytology results (most of which will be HPV positive) are referred to colposcopy and HPV-positive women without high-grade cytology are scheduled for repeat testing may be considered indicative of the value of cytology triage of HPV primary screening. In the previous co-testing section, the POBASCAM and ARTISTIC co-testing trials take this approach (see Table 81); however, these trials provide a comparison to cytology testing alone, rather than HPV testing.

HPV screening with CC triage versus CC primary screening

The Finnish RCT included was considered a fair-quality study and included women aged 25 to 65 years within the Finnish national screening program. The trial arms were HC2 testing with CC testing to triage positive HPV results, compared to CC alone. The study does not directly address the research question in the current review as the triage was using CC rather than LBC, which limits the applicability as HPV is collected into an LBC medium, but CC requires separate collection. This is unlikely to affect outcomes given the similar accuracy of the two testing strategies, unless there is a limitation in the tests being obtained from separate samples. Also, in this trial CIN1+ was treated in all but the latter years, which does not reflect Australian practice.

The Finnish trial has a threshold for referral to colposcopy of LSIL+ which is lower than in Australia. For the HPV with CC triage arm, a positive HPV test with a normal or ASC-US cytology result was followed up with repeat screening (including HPV testing) at 12 months and referral to colposcopy following either repeat ASCUS or three consecutive HPV+ test results.

The trial was limited by a high proportion of post-randomisation loss; 66,410/101,678 women (65.3%) in the HPV arm and 65,785/101,747 women (64.7%) in the control arm attended screening following randomisation. In this review data from this trial are reported by women screened rather than women randomised. Finland also has high levels of opportunistic screening.

More invasive cervical cancers were detected in the HPV arm of the trial than in the control arm among women who were screened (17 vs 9, Hazard Ratio 1.87 (0.83–4.20)). However, in women who did not attend screening, fewer invasive cervical cancers were detected in the HPV arm than the control arm (8 vs 22, HR 0.37 (0.17–0.83)). The reduction in cancer incidence in the control arm of women who did not attend screening is statistically significant and the authors have been unable to account for this difference (eg due to failed randomisation or lack of blinding) and state that it could be due to chance.

Following a single screening round, in women older than 35 years, the detection of both CIN2+ and CIN3+ was significantly increased in those screened with the HPV strategy (relative detection rate CIN2+: 1.60 (95% CI 1.34–1.90); CIN3+:1.56 (95% CI 1.19–2.05)). Relative detection rates increased to a greater degree in women younger than 35 (relative detection rate CIN2+: 1.81 (95% CI 1.45–2.26); CIN3+: 1.83 (95% CI 1.21–2.78)).

Referral to colposcopy after the first screen was low in the trial and similar across arms for women older than 35 years (HPV arm 0.92% (0.84–1.00)) vs CC arm 0.99% (0.91–1.08); P=0.18). There was also little difference for these women in numbers referred for follow-up testing, though this difference was significant (HPV triage arm 5.0% (4.8–5.2) vs CC arm 5.6% (5.5–5.8); P<0.0001). In contrast, in younger women there were more referrals to colposcopy, particularly in the intervention arm (HPV triage arm 2.6% (2.3–2.9) vs CC arm 1.9% (1.7–2.2); P<0.0001) and more follow-up testing (HPV triage arm 15.1% (14.4–15.7) vs CC arm 6% (5.6–6.5); P<0.0001) in the HPV triage arm than the control arm and overall compared to older women. The false positive rate in the HPV arm was lower than that in the cytology arm for women over 35 (false positive proportion 0.88, 95% CI 0.84–0.91), indicating a higher specificity of HPV with LBC triage than for conventional cytology. In contrast, in younger women the false positive rate was higher with the HPV-based screening strategy (false positive proportion 2.24, 95% CI 2.07–2.41).

As this trial has only one screening round, interpretation of the results are limited. Greater detection of precancerous cervical lesions may not translate to reduced cancer incidence, and reduced detection in a subsequent round would be expected. The trial demonstrates that the strategy employed does not increase further testing and colposcopy referral in women older than age 35 despite the primary screen being HPV.

HPV primary screening with LBC triage versus LBC primary screening with HPV triage

HPV FOCAL is a Canadian trial which compares HPV (HC2) primary screening with LBC triage to LBC primary screening with HPV (HC2) triage. It therefore does not directly address the research question in which the comparator is HPV primary screening without triage. The HPV screening arm has a round interval of four years at which point an exit screen is scheduled consisting of an HPV and LBC co-test. The LBC screening arm has a round interval of two years; LBC is repeated at this time and at four years an exit co-test is scheduled. Criteria for referral to colposcopy are LSIL+ which is lower than Australian practice but higher than some trials, or persistent HPV infection, or ASCUS cytology with a HPV positive test. HPV positive women who are cytology negative are scheduled for a repeat co-test at 12 months as are ASC-US positive women who are HPV negative. The trial included women aged 25 to 65.

We have rated this trial as fair quality. The key quality issues include low recruitment; letters of invitation were sent to 44,099 women but only 14,267 (32.4%) were randomised into the trial along with an additional 4,281 women identified by collaborating physicians. The trial planned to recruit 33,000 women and it is therefore currently underpowered to detect differences in the primary outcome of CIN3+ incidence. Furthermore, this may be a source of sampling bias and the randomised population may not be representative of the overall screening population. Full follow-up is also yet to be completed, Ogilvie et al. (2012) reports results for the baseline screen (including the first screening round with triage but without any follow-up testing) and the full first screening round which includes follow up tests at 12 months. The publication also only provides outcomes as rates without contingency data and does not stratify by age.

The detection of CIN2+ and CIN3+ were similar between the HPV with LBC triage testing arm and the LBC with HPV triage testing arm at the baseline screen (CIN2+: HPV arm: 9.2/1,000 (95% CI 7.4–10.9) vs LBC arm: 11.0/1,000 (95% CI 8.3–13.7) and CIN3+: HPV arm: 4.8/1,000 (95% CI 3.6–6.1) vs LBC arm: 5.0/1,000 (95% CI 3.1–6.8). However, in the overall round one

results (including follow-up testing), the detection of both CIN2+ and CIN3+ was increased in the HPV arm but remained unchanged in the LBC arm (CIN2+: HPV arm: 16.1/1,000 (95% CI 13.2–18.9) vs LBC arm: 11.0/1,000 (95% CI 8.3–13.7) and CIN3+: HPV arm: 8.0/1,000 (95% CI 5.9–10.0) vs LBC arm: 5.0/1,000 (95% CI 3.1–6.8).

Baseline colposcopy rates were similar between the two arms (HPV arm: 28.9/1,000 (95% CI 26.0–31.9) vs LBC arm: 32.1/1,000 (95% CI 27.6–36.5) but were higher in the HPV arm for round one overall (HPV arm: 57.2 /1,000 (95% CI 52.8–61.7) vs LBC arm: 33.2 /1,000 (95% CI 28.7–37.7). Although only preliminary results, the HPV FOCAL trial is showing similar results to other primary HPV trials with increased detection of CIN2+ and CIN3+ in the first round in association with increased colposcopy rates. Second round results are not yet available.

As RCT level evidence of HPV with LBC triage was identified, studies providing a lower level of evidence were excluded.

Accuracy

The AHRQ HTA (Vesco et al. 2011) did not include any accuracy studies addressing this question and this question was not specifically included in the Arbyn et al. (2012) meta-analysis. Therefore, no accuracy data are included in the evidence available for this question.

One accuracy study (Kulasingam et al. 2002) discussed in the previous HPV and cytology co-testing section (in which an abnormal result on both tests is required for a positive result) is indicative of the accuracy of a triage approach and is relevant to this research question (see Table 87). In this study the sensitivity of the co-testing strategy for CIN3+ was lower than that of HPV testing alone (women ≥30 years: HPV and cytology 49.7 (95% CI 32.9–71.5) vs HPV alone 86.0 (95% CI 59.7–96.9) while the specificity was increased (women ≥30 years: HPV and cytology 94.7 (95% CI 92.8–96.1) vs HPV alone 83.0 (95% CI 76.8–87.1). Results were similar in women aged less than 30 years (sensitivity: HPV and cytology 64.0 (95% CI 51.1–77.6) vs HPV alone 92.5 (95% CI 83.5–97.3); specificity: HPV and cytology 87.6 (95% CI 86.7–88.4) vs HPV alone 70.1 (95% 66.5–73.1)).

Prognostic value of test results

Longitudinal outcomes from RCTs and cross-sectional accuracy studies have been used to address the research questions on HPV primary screening. However, they have provided little information regarding a screening interval of five years. The trials all utilised different screening intervals, only one of which was five years (POBASCAM). Trials had the same screening interval for the intervention and the control arms with the exception of HPV FOCAL.

In an effort to address this deficiency, the systematic review is supplemented with two studies reporting on long-term cumulative CIN2+ and CIN3+ detection, stratified by test result. These studies are the extended follow-up of the ARTISTIC trial (Kitchener et al. 2011b) and the Kaiser Permanente Northwest study (Katki 2011). These studies are discussed in the AHRQ report (Vesco et al. 2011, pp. 42–44) along with two other cohorts, although the final results from the Katki study were not included in the AHRQ report as these were identified in their ancillary search. This review will focus on these two studies and the conclusions of the AHRQ and is not underpinned by a systematic review.

Extended follow-up from the ARTISTIC trial provides data from a third screening round of HPV with LBC co-testing compared with LBC alone. The rates of 6-year cumulative CIN2+ and CIN3+ detection stratified by baseline test results are reported (Table 88). Note that only 36.2%

of the original cohort was tested in round three and 22.8% of women were lost to follow-up and therefore there could be selective ascertainment bias.

Table 88 Six-year cumulative CIN2+ and CIN3+ rates associated with HPV and LBC baseline test results, women aged 20–64 years, Kitchener 2011b

Cytology baseline (ASC-US+)	HPV baseline (HC2 ≥1 RLU)	Cumulative CIN2+ (95% CI)	Cumulative CIN3+ (95% CI)
All women	-	3.88% (3.59–4.17%)	1.96% (1.76–2.17%)
Negative	-	1.41% (1.19–1.65%)	0.63% (0.48–0.80%)
Positive	-	20.53% (19.04–22.08%)	11.01% (9.87–12.23%)
-	Negative	0.87% (0.70–1.06%)	0.28% (0.18–0.40%)
-	Positive	20.12% (18.68–21.61%)	11.19% (10.05–12.40%)
Negative	Negative	0.67% (0.51–0.87%)	0.23% (0.14–0.36%)
Negative	Positive	7.73% (6.29–9.36%)	4.05% (2.98–5.36%)
Positive	Negative	3.24% (2.32–4.38%)	0.83% (0.40–1.52%)
Positive	Positive	37.44% (34.91–40.04%)	21.18% (19.03–23.45%)

The Kaiser Permanente Northern California (Berkeley) study (Katki 2011) is a prospective study of 331,818 women aged 30 years and older. The management guidelines for the cohort referred all women with LSIL+ cytology to colposcopy, women with ASCUS cytology who were also HPV positive were referred to colposcopy, and women who had ASCUS cytology and a HPV negative test were recommended for a repeat screen after one year. The screening interval was three years. Screening used conventional cytology with FocalPoint automated image analysis and HC2 HPV test.

Table 89 Cumulative five-year CIN2+, CIN3+ and ICC incidence associated with baseline HPV and cytology test results (data from Katki et al. 2011)

Cytology baseline (ASC-US+)	HPV baseline (HC2 ≥1 RLU)	Cumulative CIN2+ (95% CI)	Cumulative CIN3+ (95% CI)	Cumulative ICC (95% CI)
All women (calc)	-	0.67	0.25	0.026
Negative	-	0.96(0.73–1.26)	0.36 (0.24–0.55)	0.037 (0.016–0.087)
Positive	-	12.72 (11.06–14.61)	4.68 (3.65–5.99)	0.41 (0.16–1.07)
-	Negative	0.53 (0.40–0.70)	0.17 (0.11–0.28)	0.019 (0.008–0.046)
-	Positive	18.61 (17.03–20.31)	7.63 (6.53–8.89)	0.66 (0.37–1.16)
Negative	Negative	0.54 (0.30–0.96)	0.16 (0.06–0.39)	0.016 (0.003–0.072)
Negative	Positive	13.39 (10.03–17.77)	5.94 (3.75–9.34)	0.54 (0.20–1.49)
Positive	Negative	2.19 (1.44–3.31)	0.86 (0.42–1.73)	0.16 (0.04–0.64)
Positive	Positive	32.67 (28.48–37.30)	12.09 (9.28–15.68)	0.90 (0.14–5.67)

The two studies have broadly similar findings. Using CIN3+ as the outcome, although similar findings are observed for both CIN2+ and ICC, the risk over five years is lower for a negative HPV test than a negative cytology finding (ARTISTIC: negative HPV 0.28% (0.18–0.40) vs

negative cytology 0.63% (0.48–0.80) and Katki: negative HPV 0.17% (0.11–0.28) vs negative cytology 0.36% (0.24–0.55)). This risk does not drop further for those testing both HPV negative and cytology negative (ARTISTIC: 0.23 (0.14–0.36); Katki: 0.16 (0.06–0.39)). Similarly, in a co-testing strategy, the risk of CIN3+ is increased to a greater extent with a positive HPV test and a negative cytology finding than the converse (ARTISTIC: HPV positive/cytology negative 4.05 (2.98–5.36) vs HPV negative/cytology positive 0.83 (0.40–1.52) and Katki: HPV positive/cytology negative 5.94 (3.75–9.34) vs HPV negative/cytology positive 0.86 (0.42–1.73)).

The findings from these studies suggest that longer screening intervals up to five years may be appropriate for women who are HPV negative (Katki 2011; Kitchener 2011b). They additionally raise questions regarding the added value of cytology in HPV co-testing strategies given the greatly reduced risk of CIN3+ in HPV negative women.

Immediate versus delayed colposcopy for pLSIL/LSIL HPV positive screenees

Scenario 2 – secondary question 2.4 and Scenario 3 – secondary question 3.5

What is the comparative safety, effectiveness and cost-effectiveness of undertaking a colposcopy immediately in comparison to delaying the test in women who have pLSIL/LSIL cytology and a positive HPV test?

Summary 13: Summary of effectiveness—immediate versus delayed colposcopy

NOTE: Current Australian policy is to delay referral for colposcopy for pLSIL/LSIL for 12 months, repeat the Pap test, and if persistent (pLSIL+) refer to colposcopy. AIHW safety monitoring data show that the incidence of cervical cancer in women aged 20–69 has not changed since the introduction of this policy (AIHW 2013b).

Summary— immediate versus delayed colposcopy

The review sought to look at the effectiveness of delayed colposcopy versus immediate colposcopy in women with a positive HPV test and a pLSIL/LSIL result. For this comparison there was limited evidence.

Two included RCTs were of limited applicability so definitive conclusions cannot be made.

A modelled analysis has been conducted considering the likely benefits and harms associated with immediate versus delayed colposcopy in women HPV positive with pLSIL/LSIL cytology in the Australian setting.

Table 90 Inclusion criteria for identification of studies relevant to assessment of effectiveness of immediate versus delayed colposcopy

Selection criteria	Inclusion criteria
Population	Women with possible or low-grade squamous intra-epithelia lesion (LSIL/pSLIL) and a positive HPV result
Intervention	Immediate colposcopy
Comparator(s)	Delayed colposcopy (surveillance with repeat cytology)
Outcomes	Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC)
Reference standard	Colposcopy with biopsy for positives
Search period	From the start of databases up until 2013
Language	English
Study design [a]	RCT, pseudorandomised trial

Abbreviations: AIS = adenocarcinoma in situ; HSIL = high-grade squamous intraepithelial lesion; LBC = liquid-based cytology; SCC = squamous cell carcinoma

[a] Post-hoc inclusion criteria based upon the body of evidence identified

MSAC has not previously reviewed immediate versus delayed colposcopy in women who have pLSIL/LSIL and a positive HPV test.

A targeted review of the literature was undertaken to identify published studies that were designed to compare incidence or mortality rates of invasive cervical cancer in women with low-grade cytological abnormalities and a positive HPV test who have undergone immediate colposcopy in comparison to cytology surveillance (delayed colposcopy). The literature searches conducted to address the other questions in this review were also used to identify relevant studies.

No systematic reviews were identified that addressed the specific question as outlined in the review. However, a Cochrane protocol was identified that addressed issues of delayed versus immediate colposcopy in women with LSIL cytology (Kyrgiou et al. 2012). The first author is also the first author of a review published in 2007 addressing the same question (Kyrgiou et al. 2007). Three randomised controlled trials were included in the 2007 review. There was no requirement for studies to include women with positive HPV status. Two studies however provided some information on the relationship between HPV status and timing of colposcopy. The authors concluded that despite the limitations of the review "a general policy would be to refer all women with low grade abnormalities (ASCUS and LSIL) for immediate colposcopy. An exception could be in the management of women with ASCUS smears in countries where HPV DNA testings is available and affordable". They also stated that a possible exception was older women with LSIL.

Table 91 Relevant systematic reviews and meta-analyses of immediate versus delayed colposcopy

Study identified in review	Study title	Studies included in this review
Kyrgiou et al. 2012	Kyrgiou M, Stasinou SM, Arbyn M et al. (2012). Management of low-grade squamous intra-epithelial lesions of the uterine cervix: repeat cytology versus immediate referral to colposcopy. Cochrane Database Syst.Rev	Nil – Cochrane Protocol
Kyrgiou et al. 2007	Kyrgiou M, Stasinou SM, Arbyn M et al. (2007). Management of minor cervical cytological abnormalities: a systematic review and a meta-analysis of the literature. Cancer Treatment Reviews 33:514–520	Nil – studies did not specifically include women with positive HPV

Included studies

The inclusion criteria used for identification and inclusion of studies for the assessment of evidence for immediate versus delayed colposcopy are listed in Table 90. No studies comparing the effectiveness of immediate versus delayed colposcopy in women who have pLSIL/LSIL cytology and a positive HPV test were identified in the searches.

Two RCTs were identified that randomised women with low-grade positive cytology to either surveillance or immediate colposcopy and reported on HPV status. While a proportion of women in both of these studies had borderline cytology and a positive HPV DNA test, the majority of women were HPV negative. As such both studies are limited in their applicability to the research question. However, in the absence of studies directly applicable to this question, these studies are included and their limitations are discussed below.

A third RCT was also identified (Jakobsson et al. 2011) but was excluded from the main analysis due to applicability issues around the comparator and patient population. It is however discussed further below for completeness given the limited data on this question.

A master list of the primary studies included in the review of immediate versus delayed colposcopy effectiveness is provided in Table 92.

Table 92 Master list of trials included for evaluation of effectiveness of immediate versus delayed colposcopy in HPV women with pLSIL/LSIL

Study identified in review	Study title	Comments
HART study	Cuzick, J, Szarewski, A, Cubie, H, Hulma, G et al. (2003) Management of women who test positive for high-risk types of human papillomavirus: the HART study. Lancet 362:6	Identified in targeted literature search
TOMBOLA study	Cotton, S, Sharper, L, Little, J et al. (2010). The role of human papillomavirus testing in the management of women with low-grade abnormalities: multicentre randomised controlled trial. BJOG 117(6):645–59	Identified in targeted literature search

Study characteristics, quality and applicability

The study by Cuzick et al. (2003) was conducted in the UK as part of a cross-sectional study drawing from 161 family practices and five referral centres in the UK. Women who were HPV positive *and/or* had borderline cytology (pLSIL) were randomised to either immediate colposcopy or surveillance by HPV co-testing with cytology at 6 and 12 months, with exit colposcopy at 12 months. Randomisation was undertaken by a block randomisation method, with allocation done centrally. In the surveillance group there were 411 women, with 414 women in the immediate colposcopy group.

In terms of applicability, the patient population in the study by Cuzick et al. (2003) is different to that specified in the review question as women eligible for randomisation were those with either a postivie HPV test and/or borderline cytology. Women with a diagnosis of mild dyskaryosis (LSIL) or worse were referred for immediate colposcopy and not included as part of this study, significantly limiting the applicability of the results.

The study reported that only a small proportion of women had both borderline cytology and a positive HPV test; 9.5% (39/411) in the surveillance group and 9.4% (39/414) in the immediate colposcopy group. The majority of women were negative on cytology but with a positive HPV test; 65% in both the surveillance (267/411) and colposcopy (269/414) groups. As such the results

are more applicable to patients with a positive HPV test than they are to women with low-grade positive cytology and a positive HPV.

The publication by Cotton et al. (2010) reports on the data from the Trial of Management of Borderline and Other Low-grade Abnormal smears (TOMBOLA) to assess the value of HPV testing in triaging between cytological surveillance and colposcopy. Women aged 20–59 years were eligible for the trial if they had a low-grade abnormal cytology result (mild dyskaryosis or BNA) detected as part of the NHS cervical screening program during October 1999–October 2002. Women were randomised to either immediate colposcopy or surveillance. Randomisation was undertaken centrally; however, allocation concealment is unclear. The results of HPV testing were not disclosed to either the women or those involved in their management. HPV status however was used to stratify both randomisations. Women were classified as high-risk HPV positive if they tested positive for any of 14 high-risk HPV types, or high-risk HPV negative if they did not test positive for any of these high-risk types .Women in the surveillance group underwent cytology tests every six months, and were referred to colposcopy if any test showed moderate dyskaryosis or worse. Final follow-up was at three years, and on exit women were invited to undertake a colposcopy, with excision of the transformation zone if abnormal.

In terms of applicability, the primary research question in TOMBOLA was to assess cytological surveillance compared with immediate referral in women with low-grade cytological abnormalities. The role of HPV in this patient population was a secondary analysis and had no impact on management; therefore, the majority of the reported data do not specifically address the question posed in this review. The authors also note that the participation rate was higher in older women; one-third of the women did not attend the exit examination (Little et al. 2009). HPV status was determined using the PCR system and there is some suggestion by Arbyn et al. (2010) that HPV testing in the TOMBOLA trial has a lower sensitivity than in other trials of HPV testing.

Table 93 Characteristics and quality of studies reporting on delayed versus immediate colposcopy

Author, year Setting	Population (N, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Cuzik et al. 2003 HART UK 161 family practice clinics August 1998–November 2001	N=11,085 women N=825 (borderline cytology, HPV positive or both) Inclusion criteria Women aged 30–60 yrs attending for routine cervical screening. Exclusion criteria Abnormal smear in the past 3 years. Patient characteristics • Mean age 42 yrs	All women had HPV test and cytology done together at screening Women with borderline cytology, HPV positive or both randomised to either: Index test Cytology and HPV test (HC2) at 6-12 months followed by colposcopy if any abnormal results Comparator test Immediate colposcopy	Study design: Randomised trial within a screening setting Test threshold: Borderline cytology, positive HPV test Reference standard: Colposcopy (immediately or at 12 months) Outcomes Detection rate	Systematic review search CX P2 Q2 Screening study NHMRC level II RCT Quality: Fair Randomisation done centrally, allocation concealment unclear Applicability: Limited Women with borderline cytology rather than pLSIL/LSIL. Mixed population with HPV positive and HPV negative
Cotton et al. 2010 TOMBOLA UK – Grampain, Tayside and Nottingham region October 1999—October 2002	N=4,031 women Inclusion criteria Women aged 20–59 years who had a low-grade abnormal cytology (mild dyskaryosis or BNA) as detected in the NHS cervical screening program. Exclusion criteria Women with previous cervical treatment. Patient characteristics • 20-–29 yrs (44.2%) • 30–39 yrs (26.9%) • 40–49 yrs (20.5%) • 50–59 yrs (8.4%)	All women had HPV test and cytology done together at screening. Women were classified as HPV positive or HPV negative Women with borderline cytology, mild dyskaryosis randomised to either: Index test Repeat cytology 6 months and were referred to colposcopy if any test showed moderate dyskaryosis or worse. Comparator test Immediate colposcopy Exit of trial: colposcopy	Study design: Randomised trial within a screening setting Test threshold: Borderline cytology mild dyskaryosis Reference standard: Colposcopy and biopsy Outcomes Detection rate	Systematic review search CX P2 Q2 Screening study NHMRC level II RCT Quality: Fair Randomisation done centrally, allocation concealment unclear Applicability: Limited HPV results did not dictate management

Incidence and mortality from cervical cancer (direct evidence)

No studies have assessed the impact of delayed versus immediate colposcopy on incidence or mortality rates of invasive cervical cancer in women with a positive HPV test and positive cytology (LSIL/pLSIL).

Detection of precancerous cervical lesions (indirect evidence)

Limited data were available on the detection of CIN2+ and CIN3+ in the two included studies (see Table 94). Due to the small numbers and therefore power to detect any differences in the HART (HPV in Addition to Routine Testing) study these results are considered uninformative. No significant differences were detected in the TOMBOLA study; however, HPV status was not considered in the selection of subjects so these findings do not directly address the research question.

Table 94 Detection of CIN2 and CIN3 in studies looking at delayed versus immediate colposcopy

Study	Detection of CIN2*	Detection of CIN2*	Detection of CIN2*
-	*Threshold*	*Immediate colposcopy*	*Delayed colposcopy/surveillance*
HART	BNA	10/39 (26%)	3/39 (8%)
TOMBOLA	Mild dyskaryosis/BNA	310/880 (35%)	222/875 (25%)
-	Detection of CIN3*	Detection of CIN3*	Detection of CIN3*
-	*Threshold*	*Immediate colposcopy*	*Delayed colposcopy*
HART	BNA	4/39 (10%)	1/39 (3%)
TOMBOLA	Mild dyskaryosis/BNA	168/880 (19%)	134/875 (15%)

Abbreviations: BNA – borderline nuclear abnormality

Jakobsson 2011

As mentioned above a third RCT was also identified (Jakobsson et al. 2011) but excluded from the main analysis. This trial sought to assess whether a combination of HPV testing and cervical cytology could reduce colposcopy rates. Women were randomised to three groups including an early (2–3 months) versus delayed (5–6 months) colposcopy group – however, the difference between early and delayed group was only a matter of months therefore limiting the applicability of the study. Following randomisation, women underwent cervical cytology and HPV testing –as such it is only a small proportion of women (similar to the HART study) that is of relevance to this question.

The authors conclude that with the delayed colposcopy schedule, fewer CIN2+ lesions were detected, which would appear to support the evidence that a proportion of lesions will spontaneous resolve. However, the study population was too small to detect differences between the treatment groups.

Evidence from HPV studies

The above information also needs to be considered in the context of the HPV literature. Data from the ALTS trial are discussed in detail in the section addressing evidence for HPV triage of p/dLSIL cytology. The ALTS trial randomised women with low-grade cytological lesions to HPV triage versus repeat cytology or immediate colposcopy. This study found that fewer CIN2+ lesions were detected by repeat cytology in comparison to immediate colposcopy or HPV triage, without a change in the detection of CIN3+. These data suggested that a statistically significant proportion of prevalent CIN2 lesions in trial participants with a cytological finding of pLSIL or dLSIL regressed over two years when a strategy of repeat cytology was followed. However, these data were not stratified by those with a positive and negative HPV result, hence it is difficult to draw conclusions in relation to the current research question.

Current Australian policy

Current Australian policy is to delay referral for colposcopy for pLSIL/LSIL for 12 months, repeat the Pap test, and if persistent (pLSIL+) refer to colposcopy (NHMRC 2005). AIHW safety monitoring data indicate that the incidence of cervical cancer in women aged 20–69 has not changed since the introduction of this policy (AIHW 2013b).

Discussion

In the UK, Europe and the US, referral to immediate colposcopy is recommended in women with positive HPV results and ASCUS/LSIL cytology. This however is not without controversy as there is evidence to suggest that the majority of HPV infections will clear naturally with only a few progressing to CIN2 + or CIN3+ (Bentley 2012). The decision how to manage HPV positive women with low-grade abnormalities needs to consider the trade-off of harms and benefits.

HPV testing with partial genotyping for primary screening

Secondary research question 3.5

What is the comparative safety, effectiveness and cost-effectiveness of referring women positive for HPV16/18 +/–45 using partial genotyping systems at primary screening, immediately to colposcopy, and performing cytology triage on women positive for other oncogenic types?

Summary 14: Summary of effectiveness— HPV testing with partial genotyping

Accuracy

A subanalysis of the ATHENA trial (Castle et al. 2011) demonstrated that triage of HPV positive women with immediate referral to colposcopy for HPV 16 and/or 18 types had a similar sensitivity and PPV to triage of HPV by ASCUS+ cytology. Sensitivity was greater and PPV lower than triage by HSIL+ cytology. This does not take into account any effect of cytology triage of other HPV types, nor repeat testing and longitudinal follow-up. This study was considered a high-quality level III-2 diagnostic accuracy study.

An analysis of the intervention arm of the Swedescreen trial provided diagnostic accuracy data for alternative HPV based screening strategies. In this analysis, a theoretical strategy of referring women positive for HPV 16/18 immediately to colposcopy, and performing cytology triage on women positive for other oncogenic types had a similar sensitivity for the detection of CIN2+ and CIN3+ compared to HPV testing without triage or genotyping, with a higher positive predictive value. This strategy also had a similar sensitivity to that of HPV testing with cytology triage, but with a lower relative PPV in comparison with conventional cytology. The potential for verification bias in this analysis was considered to be high.

Prognostic value of test results

Prognostic studies demonstrate that HPV types 16 and 18 are associated with a higher risk of development of CIN2+ and CIN3+ than other oncogenic HPV types. However, the effectiveness of management strategies based upon partial genotyping of women undergoing routine screening is uncertain.

Conclusions

- HPV types 16 and 18 are associated with a higher risk of development of high-grade CIN than other oncogenic HPV types.

Table 95 Inclusion criteria for identification of studies relevant to assessment of effectiveness of LBC

Selection criteria	Inclusion criteria
Population	Women undergoing cervical cytology for the detection of cervical cancer or precancerous lesions
Intervention	HPV testing with partial genotyping for HPV 16/18 +/-45
Comparator(s)	Any
Outcomes	Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC) in women with a positive HPV result Cervical cancer incidence Mortality
Reference standard	Colposcopy with biopsy for positives
Search period	2008–2012 (HTAs); 2011–2012 (primary studies)
Language	English
Study design	Any

Further detail provided in Box 2, page 39.
Abbreviations: AIS = adenocarcinoma in situ; HPV = human papillomavirus; HSIL = high-grade squamous intraepithelial lesion; HTA = health technology assessment; LBC = liquid-based cytology; SCC = squamous cell carcinoma

The inclusion criteria used for identification and inclusion of studies for the assessment of evidence for partial HPV genotyping are listed in Table 95. The only freely available HTA or systematic review that addressed HPV genotyping in the period 2008–2012 was an Italian HTA (Ronco et al. 2012) that forms the basis of the European guidelines for quality assurance in cervical cancer screening (Arbyn et al. 2010). This HTA was partially in Italian. Sections of the report are based on a systematic review of the literature although not the section addressing HPV genotyping. In lieu of any other freely available report, the findings of this HTA are discussed below.

Therefore, the evidence presented for the consideration of the effectiveness of HPV partial genotyping here should not be considered systematic.

Table 96 Relevant HTA considered for effectiveness of HPV partial genotyping for primary cervical cancer screening

Study identified in review	Study title	Studies included in this review
Ronco et al. 2012	Ronco G, Biggeri A, Confortini M, Naldoni C, Segnan N, Sideri M, Zappa M, Zorzi M, Calvia M, Accetta G, Giordano L, Cogo C, Carozzi F, Gillio Tos A, Arbyn M, Mejier CJ, Snijders PJ, Cuzick J, Giorgi Rossi P. Health technology assessment report: HPV DNA based primary screening for cervical cancer precursors. Epidemiol Prev. 2012 May–Aug;36(3–4 Suppl 1):e1–72. Italian	Naucler et al. 2009

Incidence and mortality from cervical cancer

No studies were identified that assessed the impact of management based on HPV genotyping on incidence or mortality rates of invasive cervical cancer.

Detection of precancerous cervical lesions

The update systematic review conducted to identify HPV studies published during 2010–2012 did not identify any RCTs of HPV testing with, versus without, partial genotyping. No studies were identified that considered the accuracy of HPV subtypes 16, 18 and 45 combined for the detection of high-grade CIN.

A subanalysis of the ATHENA cohort study (Castle et al. 2011) presented accuracy data for HPV 16/18 genotyping and was included for review.

The Swedescreen RCT included for the assessment of HPV and cytology co-testing (page 163) reported the relative accuracy of detection of HPV 16/18 with cytology triage for other HPV types in comparison to conventional cytology. This study is therefore also included for the assessment of the effectiveness of HPV partial genotyping (see Table 97).

The only data on the relative accuracy of HPV genotyping included in the Italian HTA non-systematic review were also from the Swedescreen study. Other studies included in this HTA reported the risk of progression from infection by different HPV genotypes to high-grade CIN and invasive cancer (evidence from large cohort studies or nested within POBASCAM).

Studies identified in the current review reporting the prognostic risk of developing high-grade CIN based on HPV genotype were excluded from review as providing a lower level of outcomes than those from Swedescreen (Naucler et al. 2009) and the ATHENA subanalysis (Castle et al. 2011). However the general findings from these studies are considered in a non-systematic discussion below (see page 188).

Studies using HPV with partial genotyping for the triage of low-grade positive cytological lesions were also excluded. Potentially included studies that were excluded are listed in Appendix H.

A master list of the primary studies included in the review of HPV partial genotyping is provided in Table 97.

Table 97 Master list of trials included for evaluation of effectiveness of HPV partial genotyping

Study identified in review	Study title	Comments
Castle et al. 2011	Castle PE, Stoler MH, Wright TC et al. (2011). Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: A subanalysis of the ATHENA study. Lancet Oncol 12, 880–890	Identified in updated literature search
Swedescreen	Naucler P, Ryd W, Tornberg S et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. J Natl Cancer Inst. 2009;101:88–99	Included in Vesco et al. 2011 & Ronco et al. 2012

Study characteristics, quality and applicability

The subanalysis of the ATHENA study (Castle et al. 2011) is a high-quality diagnostic accuracy study of approximately 40,000 women aged 25 and over. One per cent of women had received a HPV vaccine. The study reported accuracy data for the triage of HPV positive women with HPV 16, 18 or both in comparison with triage by ASCUS+ cytology for the detection of CIN3+ and CIN2+. The study used the cobas HPV test. Women who were HPV positive or had ASCUS+ cytology were referred to colposcopy. A random sample of women negative on both cytology and HPV testing also underwent colposcopy, however only crude accuracy data were reported for the partial genotyping outcomes (ie, accuracy data adjusted for verification were not reported). The

reported sensitivity is thus considered invalid as an absolute accuracy measure and has been recalculated by the reviewers as a relative sensitivity measure using the total number of women tested (40, 901) as the denominator. Colposcopists and pathologists were blinded to HPV and cytology results. The characteristics of this study are summarised below in Table 98.

The characteristics of the Swedescreen RCT are summarised in Table 81 above, in the section considering HPV testing for primary screening and also Table 98 below, specifically with regard to this analysis. Briefly, Swedescreen is a RCT of 12, 527 women receiving HPV (with PCR) co-testing with conventional cytology or conventional cytology alone. Women were referred for colposcopy based upon cytological high-grade lesions. Women testing positive for HPV without high-grade cytology were referred for repeat HPV testing at 12 months; those with persistent HPV were then referred to colposcopy. As such, the screening strategy for women with low-grade cytological lesions is similar to that of a HPV triage strategy. The age range for inclusion was narrow (32 to 38 years) and the study was conducted over two screening rounds, with a mean of 4.1 years of follow-up.

The study by Naucler et al. (2009) is an analysis of the intervention arm of the Swedescreen trial, considering multiple different HPV-based screening strategies. As actual colposcopy referral and verification was based on the trial protocol, rather than the proposed strategies, only women with cytological high-grade lesions (at four sites, ASCUS at one site) or *persistent* HPV underwent colposcopy. The authors conducted sensitivity analyses to account for non-attending women and for women with inadequate or missed screening tests. However, the bias from lack of immediate verification in women who were HPV positive and cytologically low grade or negative cannot be accounted for. Therefore, the potential for verification bias is considered high and the study is considered low quality for this analysis. The Italian HTA stated that "it must be kept in mind that, as HPV positive women in fact underwent 'cytological triage', it was impossible to observe gains in sensitivity in comparison to [screening with conventional cytology]." The data presented here for relative detection rates of HPV genotyping from the Swedescreen trial (Naucler et al. 2009) are considered as a paired diagnostic accuracy study (NHMRC level of evidence III-2), as the data are an analysis of the intervention arm of the trial only.

Table 98 Characteristics and quality of ATHENA study reporting on the relative accuracy of HPV partial genotyping

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality & applicability
Castle et al. 2011 ATHENA study US 61 sites May 2008–August 2009	N=40,901 Inclusion criteria • ≥25 years • Routine screening population • Intact uterus • Willing to undergo colposcopy with biopsy Exclusion criteria • Pregnant • Treatment CIN within 12 months • In HPV treatment trial Patient characteristics • Mean age 41.9 yrs +/– 11.3 (SD) • Prevalence CIN2+ 1.1%, CIN3+ 0.7%	Index test HPV genotyping as triage of HPV positive (HPV 16,18 by cobas HPV test) Comparator test Cytology triage of HPV positive	Study design: Cohort study Test threshold: Cytology: ASCUS+ Reference standard: CIN3+, CIN2+ Reference standard: Colposcopy with biopsy within 12 weeks Referral threshold: ASCUS+ or HPV+, plus random sample of negatives Outcomes Detection of biopsy-confirmed CIN3+, CIN2+	Systematic review C1 P2 Q1 NHMRC level III-2 Quality: High (diagnostic accuracy) Prospective study with valid reference standard and clear outcomes reporting Applicability: Limited Applicable index test & comparator Non-vaccinated population No cytology triage of other HPV types
Naucler et al. 2009 Swedescreen study Sweden 5 cities May 1997–November 2000	N=6,257 Inclusion criteria • 32–38 years • Screening population • Histologically verified CIN2+ within 6 months of enrolment or colposcopy Exclusion criteria • None Patient characteristics	Index test HPV DNA (PCR) co-testing, with 14 HPV type-specific probes performed post-hoc on frozen samples (from split-samples). Comparator test Conventional cytology	Study design: Diagnostic accuracy study (within intervention arm of RCT) Test threshold: Cytology: ASCUS+ Reference standard: CIN3+, CIN2+ Reference standard: Colposcopy ±biopsy within 6 months Referral threshold high-grade cytology or persistent HPV at 12 months When no colposcopy performed result classified as no high-grade CIN	Systematic review C1 P2 Q1 NHMRC level III-2 Quality: Low (diagnostic accuracy) Post-hoc analysis, reference standard not performed on all positive patients as cytology triage conducted within main trial. Applicability: Limited Applicable index test & comparator

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality & applicability
	• Mean age 35.1 yrs • Prevalence CIN2+ 1.4%, CIN3+ 0.8%		Outcomes Sensitivity, specificity, PPV, NPV (absolute [invalid] & relative to CC) for CIN2+, CIN3+. Number of screening tests to detect one case CIN2+, CIN3+	Non-vaccinated population Narrow age range

Abbreviations: ASCUS = atypical squamous cell, undetermined significance; ATHENA = Addressing THE Need for Advanced HPV diagnostics; CIN = cervical intraepithelial neoplasia; HPV = human papilloma virus

Accuracy

Castle et al. (2011) demonstrated that triage of HPV positive women by genotyping of HPV 16, 18 or both had a similar sensitivity and PPV to triage by ASCUS+ cytology for the immediate detection of CIN3+ or CIN2+ (Table 99). This strategy was more sensitive, but had a lower PPV than triage with LSIL+ or HSIL+ cytology (Table 99). This comparison, however, does not consider any additional value of cytology triage for other oncogenic HPV types. Nor does it consider additional detection or regression of lesions that may occur over time with repeat testing or delayed colposcopy.

Table 99 Accuracy of triage strategies for HPV positive women by HPV 16, 18 genotyping or cytology triage relative to all HPV+ referred to colposcopy

	Triage strategy	Relative sensitivity (95% CI)	PPV (%) (95% CI)
Study	Detection of CIN2+	Detection of CIN2+	Detection of CIN2+
ATHENA (Castle et al. 2011)	No triage (all HPV+ to colp)	-	10.9 (10.8–10.9)
-	ASCUS+	0.53 (0.44-0.62)	21.3 (19.4–23.3)
-	LSIL+	0.39 (0.32-0.47)	26.4 (23.5–29.5)**
-	HSIL+	0.20 (0.16-0.26)	58.8 (50.6–66.5)**
-	HPV 16	0.44 (0.37–0.53)	24.2 (21.8–26.9)
-	HPV 16 or 18	0.52 (0.44-0.62)	20.4 (18.6–22.3)
Study	Detection of CIN3+	Detection of CIN3+	Detection of CIN3+
ATHENA (Castle et al. 2011)	No triage (all HPV+ to colp)	-	7.2 (7.2–7.2)
-	ASCUS+	0.52 (0.44-0.62)	14.1 (12.6–15.8)
-	LSIL+	0.53 (0.43-0.65)	17.9 (15.5–20.6)**
-	HSIL+	0.40 (0.32-0.50)	50.4 (42.5–58.2)**
-	HPV 16	0.50 (0.41-0.62)	18.3 (16.3–20.6)**
-	HPV 16 or 18	0.60 (0.49-0.73)	15.5 (14.0–17.1)

Abbreviations: HPV = Human Papillomavirus, PPV = positive predictive value

In the Swedescreen analysis, screening for HPV 16/18 and performing cytological triage for other HPV types (with repeat HPV testing for those with normal cytology) detected the same number of CIN2+ and CIN3+ lesions as HPV testing with no further triage, or HPV testing with triage by cytology (Table 101). This strategy demonstrated a slightly higher PPV than HPV testing without further triage (relative PPV compared to conventional cytology for CIN2+ = 0.66, 95% CI = 0.51–0.86; for HPV testing alone compared to CC = 0.45, 95% CI 0.34–0.59). The PPV was still somewhat lower than that for triage of HPV screening with cytology (relative PPV for CIN2+ compared to conventional cytology screening = 0.90, 95% CI 0.70–1.16). However, as referral to colposcopy in the intervention arm of the trial (ie for all strategies analysed) was based on high-grade cytology or persistent HPV, these data are prone to verification bias. It is presumed that the increased sensitivity observed over that of cytology screening alone is due to persistent HPV 16/18 infections. Whether any additional high-grade CIN detected by the HPV-based screening strategies is destined to progress or regress is unknown.

In contrast, screening for HPV 16/18 alone without triage of other HPV types had a worse relative accuracy profile than either conventional cytology, HPV testing alone, or HPV testing with cytology triage (Table 101). The authors of the 2012 Italian HTA stated that there was no evidence

to recommend restricting direct referral to colposcopy to women with infection by HPV types 16 or 18 at that time.

The Naucler et al. (2009) study also reported the number of screening tests performed to detect one case of CIN3+. For conventional cytology this was 169.1, HPV testing alone 130.4, HPV 16/18 genotypes 223.5, and HPV 16/18 with cytology triage of other HPV types 140.0. Again these data are prone to verification bias.

Table 100 Relative accuracy of screening and triage strategies in comparison to conventional cytology screening (adapted from Naucler et al. 2009)

	Screening strategy	Relative sensitivity (95% CI)	Relative PPV (95% CI)
Study	Detection of CIN2+	Detection of CIN2+	Detection of CIN2+
Swedescreen (Naucler et al. 2009)	HPV test	1.34 (1.16–1.54)	0.45 (0.34–0.59)
-	HPV test with cytology triage	1.34 (1.16–1.54)	0.90 (0.70–1.16)
-	HPV 16 or 18	0.71 (0.55–0.91)	0.59 (0.43–0.81)
-	HPV 16 or 18, cytology triage of other HPV+	1.34 (1.16–1.54)	0.66 (0.51–0.86)
Study	Detection of CIN3+	Detection of CIN3+	Detection of CIN3+
Swedescreen (Naucler et al. 2009)	HPV test	1.30 (1.09–1.54)	0.44 (0.30–0.64)
-	HPV test with cytology triage	1.30 (1.09–1.54)	0.87 (0.60–1.26)
-	HPV 16 or 18	0.76 (0.56–1.02)	0.63 (0.41–0.98)
-	HPV 16 or 18, cytology triage of other HPV+	1.30 (1.09–1.54)	0.64 (0.44–0.94)

Abbreviations: HPV = Human Papillomavirus, PPV = positive predictive value

Prognostic value of test results

The Italian HTA included discussion of the incidence of high-grade CIN stratified by HPV genotype from two cohort studies with longitudinal follow-up: the Portland cohort (Kahn et al. 2005) and the Guanacaste cohort (Castle et al. 2009). This was not based on a systematic review. The following section considers this discussion plus more recent data from the Portland cohort (Schiffman et al. 2011) and a Danish cohort study of persistent HPV infection with 13 years of follow-up (Kjaer et al. 2010). Cumulative detection rates of high-grade CIN from the ARTISTIC trial (Kitchener et al. 2011b) are also presented. This discussion should not be considered systematic.

The Kaiser Permanente Portland (USA) cohort (Khan 2005; Schiffman et al. 2011) is a prospective study of approximately 20,000 women recruited between 1989 and 1990 with up to 18 years of follow-up. The 10-year cumulative incidence of CIN3+ after HPV detection at recruitment for women aged ≥30 years who were cytologically normal at recruitment was 20.7% (95% CI 8.6–32.8 for HPV 16; 17.7% (95% CI 0.0–36.0) for HPV 18; and 1.5% (95% CI 0.3–2.7) in women positive for other HPV types using the HC2 test (Khan et al. 2005). The risk in women who were HPV negative at recruitment was 0.5% (95% CI 0.3–0.7).

A Guanacaste (Costa Rica) cohort of 2,282 women enrolled between 1993 and 1994 reported on the cumulative incidence of CIN2+ and CIN3+ following persistent HPV infection in women after about one year (9–21 months) (Castle et al. 2009). The five-year cumulative incidence of CIN3+ in women aged ≥30 years regardless of cytology at recruitment was 24.7% (95% CI 8.8–40.5) for HPV 16; 9.1% (95% CI −7.9–26.1%) for HPV 18 and 19.8% (95% CI 5.0– 34.9%) in

women HPV positive for other carcinogenic types. The rate in HPV negative women was 0.26% (95% CI −0.03–0.56%)

A Danish cohort study of 8,656 women enrolled between 1991 and 1993 reported on the cumulative incidence of CIN3+ in women with normal cytology at baseline (Kjaer et al. 2010). The women were aged 20–29 years at enrolment and after 12 years of follow-up, the risk of developing CIN3+ was 26.7% (95% CI 21.1–31.8%) for HPV 16 at baseline and 19.1% (95% CI 10.4–27.3%) for HPV 18 at baseline. Women who had persistent HPV 16 infection for two years (two positive tests) had a risk of developing CIN3+ at 3, 5, and 12 years of follow-up of 8.9% (95% CI 2.5–14.9%), 23.8% (95% CI 14.1– 32.4%), and 47.4% (95% CI 34.9–57.5%) respectively. Over the full 13.4 years of follow-up, the risk of developing CIN3+ was 26.0% for HPV 16, 15.4% for HPV 18 and 6.4% for HPV 45.

The ARTISTIC trial (Kitchener et al. 2011b) also included HPV genotyping over three screening rounds. The characteristics of this RCT are presented and discussed above (see Table 55). Data on cumulative CIN2+ and CIN3+ detection over three screening rounds by HPV type at baseline are reproduced from Kitchener et al. (2011b) below (Table 101). The authors state that: "The dominant effect of type 16 or 16/18 combined is clearly evident, particularly for CIN3+ with a 3-fold greater rate for type 16 (30.35%) compared with types 31, 33, 45, 52 and 58 combined (10.68%)."

Table 101 Six-year cumulative CIN2+ and CIN3+ detection rates associated with HPV genotypes at baseline, women aged 20–64 years, ARTISTIC trial, Kitchener et al. 2011b

HPV at baseline	N	Cumulative CIN2+ (%) (95% CI)	Cumulative CIN3+ (%) (95% CI)
Negative	20,697	0.87 (0.70—1.06)	0.28 (0.18—0.40)
Positive	3,813	20.12 (18.68—21.61)	11.19 (10.05—12.40)
16	827	43.55 (39.75—47.45)	30.35 (26.70—34.24)
16/18	1,098	38.49 (34.25—41.81)	25.91 (22.90—29.09)
31/33/45/52/58	895	23.93 (20.82—27.26)	10.68 (8.46—13.23)
Other	1,820	6.95 (5.61—8.50)	2.73 (1.86—3.85)

Self-collection of HPV testing for primary screening of underscreened women

Secondary research question 3.4

What is the comparative safety, effectiveness and cost-effectiveness of including self-collected samples for HPV testing for underscreened and unscreened women to supplement the organised screening program using practitioner-collected HPV samples, compared with the existing protocol?

Summary 15: Summary of effectiveness—HPV self-collection

Accuracy

HPV self-collection showed moderate-high sensitivity and comparably high specificity for detecting CIN2+ compared to clinic HPV testing in 9 of 10 studies identified, with a relative sensitivity of 0.62–1.00 and relative specificity of 0.93–1.00. The evidence available from these studies suggests self-collection HPV accuracy varies for different types of sampling devices and HPV tests (Snijders et al. 2013).

Screening participation

There is strong consistent evidence that providing HPV self-collection kits to women who do not attend for cervical screening or are under-attenders improves screening participation (Snijders et al. 2013).

The size of improved participation rates achieved varies across different populations. Eight RCTs have reported improvements in participation rates of between 4% and 24% compared to an additional recall letter.

Factors shown to have an impact on participation rates are: whether the self-collection kit is mailed with an invitation to screen or available on request, ethnicity, and screening history with one trial showing that women who were unscreened are more likely to respond than women who were not screened in accordance with the recommended screening interval.

Three studies that have examined adherence to screening follow-up for this population have reported high adherence (68–100%) to follow-up after a positive HPV test and very high adherence to colposcopy referral.

Conclusions

- HPV self-collection has a moderate-high sensitivity and comparably high specificity for detecting CIN2+ compared to clinic HPV testing.
- The accuracy of HPV self-collection varies for different types of sampling devices and HPV tests.
- Providing HPV self-collection kits to women who do not attend for cervical screening or who are under-attenders improves screening participation. The magnitude of this is uncertain.

Table 102 Inclusion criteria for identification of studies relevant to assessment of the safety and effectiveness of HPV self-collection

Selection criteria	Inclusion criteria
Population	Women eligible for cervical screening who are unscreened or underscreened
Intervention	HPV self-collection
Comparator(s)	Cytology testing
Outcomes	Accuracy (relative or absolute sensitivity and specificity) for the detection of precancerous high-grade cervical lesions (CIN2+, CIN3+, AIS, SCC) Screening participation rates Cervical cancer incidence Mortality
Search period	2008–2012 (HTAs); 2011–2012 (primary studies)
Language	English
Study design	RCTs or high-quality observational studies

Included studies

No studies have reported on the impact of self-collection for HPV testing on cervical cancer incidence or mortality. Therefore, this review reports on evidence about the accuracy of self-collection for HPV testing versus practitioner collected HPV testing for the detection of CIN2+ and the impact of self-HPV testing on screening participation rates. The inclusion criteria used for identification and inclusion of studies for the assessment of evidence for HPV self-collection are listed in Table 102.

From the HTA search, the Peirson et al. (2002) review was the most recent HTA report to investigate the effectiveness of self-collection for HPV testing. This review included an assessment of self-collection for HPV testing acceptability and impact on participation rates with a literature search conducted from 2005 to February 2011. It was classified as a high-quality systematic review. The evidence for participation rates in screening non-attenders applicable to the present review consisted of one RCT (Gök et al. 2010) and one uncontrolled study (Sanner et al. 2009).

The literature search conducted to update the HTA search identified one systematic review that addressed questions about both accuracy and participation rates applicable to the present review (Snijders et al. 2013). This review included the two eligible studies identified from the Peirson et al. (2012) review.

One additional study published after Snijders et al. review was also identified as eligible for inclusion (Tamalet et al. 2013).

Study characteristics, quality and applicability

Systematic review

The Snijders et al. (2013) review is included as the most recent comprehensive review of the accuracy of self-collection for HPV testing and impact on screening participation rates.

It was designed to assess the accuracy of self-collection for HPV testing versus cytology or HPV performed by a clinician, and impact on participation rates with a literature search conducted up to January 2012. The authors did not report information about appraisal of included studies for classification of review quality, but the review met all other quality criteria.

The evidence for the accuracy of self-collection for HPV testing included 10 studies that compared self-collection with clinic-based collection in a screening setting which are applicable to the present review.

The evidence for participation rates included eight RCTs, one cohort study and one uncontrolled study conducted in screening non-attenders that are applicable to the present review.

The review questions and methods are summarised in Table 103.

Table 103 Systematic reviews examining accuracy and participation rates for HPV self-testing: methods and quality appraisal

Author, year	Review questions	Search methods	Evidence	Quality
Snijders 2013	Is the accuracy of self-HPV comparable to clinic-cytology for detection of CIN2+? Is the accuracy of self-HPV comparable to clinic-HPV for detection of CIN2+? Does self-HPV increase screening attendance compared to clinic sampling?	*Search period:* January 1992–January 2012 *Sources:* PubMed *Included studies:* Any study design that addressed the PICO-defined research questions Any language	*Accuracy* Self-HPV vs cytology (10 studies) o screening setting (5 studies) o referral setting (5 studies) Self-HPV vs clinic-HPV (21 studies) o screening setting (10 studies) o referral setting (11studies) *Attendance* 9 studies in non-attenders (7 RCT, 1 cohort study, 1 uncontrolled study)	Quality: unclear Appraisal of included studies not explicit All other quality criteria met

Abbreviations: RCT = randomised controlled trial, PICO = Population, Intervention, Compartor, Outcomes, self-HPV = self collection for HPV testing

Primary study

One additional randomised controlled trial is included as an update to the Snijders et al (2012) review. This trial enrolled 9334 women aged 25-69 years living in France who had not had a cytology screening test recorded for more than two years and did not respond to an initial mailout invitation for a free cytology test at a local clinic (Tamalet et al. 2013). Women were randomised to receive a second invitation for cytology or a home self-HPV. It is classified as providing NHMRC Level II high-quality evidence and reports outcomes for a population applicable to the present review.

Is it accurate?

The Snijders et al. (2013) review reports on evidence from 10 studies that compared the accuracy of self-collected versus clinic-based collection for HPV testing in screening settings (over 35,000 subjects). All studies used colposcopy as the reference standard for all women (four studies) or all women with abnormal HPV and/or cytology results (six studies).

Self-collection HPV devices included brushes (five studies) and swabs (five studies). The characteristics of these studies are presented in Table 104.

Table 104 Characteristics of primary studies reporting on the accuracy of HPV self-collection

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Wright et al. 2000 South Africa (outpatient clinics in peri-urban settlement, outside Cape Town) Recruitment period: January 1998–April 1999	N=1,415 women (35–65 years old) Inclusion criteria • Previously unscreened South African black women aged 35–65, who were participating in an ongoing cervical cancer screening study Exclusion criteria • Not stated	Index test: HPV testing (using HC2 HPV DNA assay) of self-collected vaginal swab Comparator test: HPV testing of a clinician-obtained cervical sample (also: Pap test, VIA, cervicography)	cross-sectional accuracy study Reference standard: colposcopy (all women with abnormal results on any of the screening tests)	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Belinson et al. 2001 Shanxi Province, China Recruitment period: not stated	N=1,997 women (aged 35–45) Inclusion criteria • Non-pregnant women • No history of cervical screening, pelvic radiation, or hysterectomy Exclusion criteria • Not stated	Index test: Self-test for HR HPV using a Dacron swab (other tests also used as part of the study: fluorescence spectroscopy, liquid-based cytology, HPV testing, VIA, colposcopy) Comparator test: n/a (all patients received identical screening tests)	cross-sectional accuracy study Reference standard: Colposcopy (all women)	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Lorenzato et al. 2002 Recife, Brazil Recruitment period: not stated	N=253 women (16–88 years old) Inclusion criteria • Any woman who came for screening at the Institute of Mother and Child Health in Pernambuco and the Cancer Hospital in Pernambuco, including pregnant women, without age restriction Exclusion criteria • Women without a cervix and women who were too ill or unwilling to participate in the study	Index test: MY 09/11 PCR test (self-collection) Comparator test: Colposcopists-collected samples, using Ayre's spatula for ectocervix and cytobrush endocervical brush for the endocervical canal	cross-sectional accuracy study Reference standard: Colposcopy (all women)	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Belinson et al. 2003 Shanxi Province, China Recruitment period: not stated	N=8,497women (27–56 years old) Inclusion criteria • Non-pregnant women, with no history of pelvic radiation or hysterectomy Exclusion criteria • NR	Index test: Self-test using a conical-shaped brush, placed high in the vagina, and rotated 3-4 times. Tested with the HC II HPV test Comparator test: HPV test on a physician-obtained sample (HC II), and liquid-based cervical cytology (AutoCyte)	cross-sectional accuracy study Reference standard: Colposcopy (all women with abnormal HPV test result and/or abnormal cytology result)	From Snijders et al. 2013 Diagnostic accuracy study level III-2

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Salmeron et al. 2003 Morelos, Mexico Recruitment period: May–October 1999	N=7,732 (15–85 years old) Inclusion criteria • Women attending CC screening services at any one of the 23 health units that make up the Morelos Cervical Cancer Screening Programme, between the ages of 25–65 Exclusion criteria • Women with history of CIN2+ or cervical cancer, with a previous hysterectomy, or who were pregnant	Index test: Self-sample, using a Dacron swab, tested with HC II HPV test Comparator test: Pap test, HC II HPV test using clinician-collected sample (conical cytobrush)	cross-sectional accuracy study Reference standard: Colposcopy (all women with ASCUS+ cytology, and/or abnormal self-sample HPV results, and/or abnormal clinician-collected sample HPV results)	From Snijders et al.2013 Diagnostic accuracy study level III-2
Holanda et al. 2006 Brazil (rural districts of the cities of: Crato, Sobral, Pedra Branca, Redencao, Ilbiapina) Recruitment period: August–December 2002	N=878 women (15–69 years old) Inclusion criteria • Sexually active women 15–70 years old, living in the specified rural districts Exclusion criteria • Pregnant women, women who have had hysterectomies. (In case of menstruation, sexual intercourse within the preceding 24 hrs, use of vaginal creams, a later home visit for sample collection was scheduled)	Index test: Self-sample, using a vaginal brush, tested with HC II HPV test Comparator test: Gynaecologist-collected sample, using an Ayre spatula, tested with HC II HPV test	cross-sectional accuracy study Reference standard: Colposcopy (all women), biopsy when applicable	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Szarewski et al. 2007 London (Margaret Pyke Centre and Hounslow Primary Care Trust) Recruitment period: January 2001 to November 2004	N=925 women (20–early 60s years old) Inclusion criteria • Women who were due for a routine screening smear, and who had not previously had ablative or excisional treatment of the cervix Exclusion criteria • Not stated	Index test: Self-sample, using a cotton swab (Digene kit) Comparator test: Doctor- or nurse-collected cervical smear, and a clinician HPV test using Digene Hybrid Capture II kit	cross-sectional accuracy study Reference standard: Women with either an abnormal cervical smear or positive HPV test result were offered colposcopy, with biopsy. Randomly selected 5% of women who tested negative on all tests were asked to attend for colposcopy	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Qiao et al. 2008 Shanxi province, China Recruitment period: May–June 2007	N=2,530 women (30–54 years old) Inclusion criteria • Aged 30–54, married, not pregnant, no history of cervical cancer, not menstruating, no hysterectomy, no history of pelvic radiation Exclusion criteria • Not stated	Index test: careHPV test of self-obtained vaginal specimen Comparator test: careHPV test by provider-obtained cervical specimen. (Also: VIA, HC2, LBC tests.)	Cross-sectional accuracy study Reference standard: Colposcopy (all women) with directed biopsy and endocervical curettage as required	From Snijders et al. 2013 Diagnostic accuracy study level III-2

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Belinson et al. 2010 Shanxi Province, China (5 sites) Recruitment period: May 2006–April 2007	N=2,625 women (16–54 years old) Inclusion criteria • Not pregnant, no history of pelvic radiation, hysterectomy, previous treatment for cervical cancer or seropositivity for HIV Exclusion criteria • Not stated	Index test: Self-collected vaginal specimen, using conical shaped brush, tested with HC II Comparator test: Physician-collected endocervical samples tested with HC II	cross-sectional accuracy study Reference standard: Self-collected or physician-collected endocervical samples testing positive for HPV by HC II, or cervical cytology other than normal or ASCUS, underwent colposcopy (using Preventive Oncology International microbiopsy protocol)	From Snijders et al. 2013 Diagnostic accuracy study level III-2
Belinson et al. 2011 Guangdong Province, China (7 sites) Recruitment period: April 2009–April 2010	N=10,000 women (25–59 years old) Inclusion criteria • 25–59 years old, not pregnant, no cervical cancer screening for at least 3 years, no prior hysterectomy, no prior pelvic radiation Exclusion criteria • Not stated	Index test: Self-collected vaginal sample, using either a novel collection device (POI/NIH self-sampler) or conical-shaped brush. Tested using Cervista and MALDI-TOF. Comparator test: Clinician collected endocervical specimens (tested using Cervista and MALDI-TOF). Also, cervical cytology	cross-sectional accuracy study Reference standard: Women testing positive on self- or clinician-collected sample, or had cervical cytology >ASCUS were asked to return for colposcopy (using the Preventive Oncology International microbiopsy protocol)	From Snijders et al. 2013 Diagnostic accuracy study level III-2

Abbreviations: NR – not reported, VIA -Visual Inspection with Acetic Acid, POI/NIH - the Preventive Oncology International/National Institutes of Health self-sampler, MALDI-TOF - mass array matrix-assisted laser desorption/ionization time-of-flight

Estimates of the sensitivity of self-collection for HPV testing varied widely from 47% to 94% across the 10 studies. The relative sensitivity of self-collection versus clinic-based collection ranged from 0.62 to 1.00, corresponding to an absolute difference in sensitivity between self-collection and clinic-based collection of 0–29% (Table 105).

Estimates of the specificity of self-collection for HPV testing varied from 76% to 100%. The relative sensitivity of self-collection versus clinic-based collection ranged from 0.93 to 1.00, corresponding to an absolute difference in sensitivity of 1–5% (Table 105).

Snijders et al. (2013) noted that the use of different collection devices and HPV tests and protocols may have contributed to the variations in the accuracy of self-collection for HPV testing across studies. Across their broader review, they observed studies that used brushes and the lavage-based Delphi screener as self-HPV devices with clinically validated HPV tests performed well, whereas some studies using swabs reported poorer performance. One study that used two different types of HPV tests also provided evidence that indicates the type of HPV test also has an impact on the accuracy of self-collection (Belinson et al. 2011).

Table 105 Accuracy self-collection HPV versus clinic-collected HPV for detection of CIN2+ in screening population (from Snijders et al. 2013)

Study	N	Self-HPV Sens. 95% CI	Self-HPV Spec. 95% CI	Clinic-HPV Sens. 95% CI	Clinic-HPV Spec. 95% CI	p-value	Relative accuracy Sens.	Relative accuracy Spec.
Wright et al. 2000	1,365	66	81	84	83	0.01	0.79	0.98
Belinson et al. 2001	1,997	83	86	95	85	NR	0.87	1.01
Lorenzato et al. 2002	253	47	86	76	88	<0.03	0.62	0.98
Belinson et al. 2003	8,497	88	77	97	80	<0.001	0.91	0.96
Salmeron et al. 2003	7,732	71	89	93	92	-	0.76	0.97
Holanda et al. 2006	878	88.9	66.7	88.9	72.0	NS	1.00	0.93
Szarewski et al. 2007	920	81	82	100	85	NS	0.81	0.96
Qiao et al. 2008	2,388	72.9	87.7	84.3	87.5	0.06	0.86	1.00
Belinson et al. 2010	2,625	80.9	88.6	97.9	90.2	0.008	0.83	0.98
Belinson et al. 2011								
Cervista	8,556	70.9	86.1	95.0	90.3	0.001	0.75	0.95
MALDI-TOF	-	94.3	87.6	94.3	89.4	NS	1.00	0.98

Abbreviations: NS = not statistically significant, NR = not reported, sens = sensitivity, spec = specificity

In addition, this review reports on evidence from five studies that compared the accuracy of self-collection for HPV testing versus clinic-based cytology for detecting CIN2+ in screening settings. Based on this evidence, the reviewers concluded that self-collection for HPV testing is at least as sensitive as clinic-based cytology and more sensitive in some studies (sensitivity 66–94%), although generally less specific (specificity 51–94%)(Snijders et al. 2013).

Does it change screening participation?

Ten studies investigating the impact of self-collection for HPV testing on screening participation among non-attenders or under-attenders have enrolled over 91,000 women. All studies have been conducted in developed countries in Europe. These studies include eight RCTs and one controlled trial that have compared self-collection for HPV testing with a repeat recall letter to report on participation rates; and one uncontrolled study. The characteristics of these studies are presented in Table 106.

Table 106 Characteristics of primary studies reporting on participation using self-HPV testing in underscreened and unscreened women

Author, year Setting	Population (*N*, inclusion criteria)	Test comparison/s	Study design	Source Quality and applicability
Bais et al. 2007 Netherlands, urban setting	N=2,830 women (30–50 years old) Inclusion criteria: Women who did not respond to two invitations for screening	Index test: Mailed HPV self-test (PCR genotyping with cervicovaginal brush sampling) Comparator test: Pap invitation letter	Pseudo-randomised trial	Snijders et al. 2013 NHMRC Level III-1
Sanner et al. 2009 Sweden, urban setting	N=2,829 (30–58 years old) Inclusion criteria: Women who had not attended the organised screening for > or = 6 years	Index test: Mailed invitation to order a HPV self-test (HC2 with swab sampling device)	Uncontrolled single arm study	Snijders et al. 2013 NHMRC level IV
Gök et al. 2010 Netherlands, urban setting	N=28,073 women (30–60 years old) Inclusion criteria: Women who had not had a Pap test in 5 years and did not respond to 1 invitation for screening	Index test: Mailed HPV self-test (HC2 assay with lavage sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Gök et al. 2012 Netherlands, urban setting	N=26,409 women (30–60 years old) Inclusion criteria: Women who had not attended cervical cancer screening in the last year after a reminder invitation for screening	Index test: Mailed HPV self-test (HC2 assay with cervicovaginal brush sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Virtanen et al. 2011a Finland, urban setting	N=4,160 (30–60 years old) Inclusion criteria: Women who did not respond to one regular invitation for screening	Index test: Mailed HPV self-test (HC2 assay with lavage sampling device Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Virtanen et al. 2011b Finland, urban setting	N=8,699 women (30–60 years old) Inclusion criteria: Women who did not respond to two invitations for screening	Index test: Mailed HPV self-test (HC2 assay with lavage sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Szarewski et al. 2011 United Kingdom, urban setting	N=3,000 (25–64 years old) Inclusion criteria: Women who did not respond to two invitations for screening	Index test: Mailed HPV self-test (HC2 assay with swab sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Giorgi et al. 2011 Italy, urban and rural settings	N=1,235 (35–65 years old) Inclusion criteria: Women who did not respond to one regular invitation for screening	Index test: Mailed HPV self-test (HC2 assay with lavage sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Wikstrom et al. 2011 Sweden, urban setting	N=4,060 (39–60 years old) Inclusion criteria: Women who had not participated in screening for >6 years	Index test: Mailed HPV self-test (HC2 assay with swab sampling device) Comparator test: Pap invitation letter	RCT	Snijders et al. 2013 NHMRC Level II
Tamalet et al. 2013 France, urban setting	N=9,334 (25–69 years old) Inclusion criteria: Women who had not had a Pap test for ≥2 years	Index test: Mailed HPV self-test (PCR assay with swab sampling device) Comparator test: Pap invitation letter	RCT	Systematic review update search NHMRC Level II

Overall, the studies report self-collection for HPV testing screening participation rates of between 8.7% and 39% in non- or under-attenders (Table 107). The lowest participation rate (8.7%) was reported for non-attenders who were sent a self-collection HPV test on demand (Giorgi-Rossi et al. 2011). This trial included three arms: self-collection for HPV testing on demand, self-collection for HPV testing mail-out and repeat recall. Participation rates for repeat recall to cytology or HPV testing (13.9% and 14.9% respectively) were higher than achieved in the self-collection on demand group. In contrast, a participation rate of 19.6% was achieved for the group of women who were directly sent a self-collection kit to complete and return by mail (Giorgi-Rossi et al. 2011).

Table 107 Participation rates for self-HPV versus recall of unscreened and underscreened women (from Snijders et al. 2013)

Study	N	Self-HPV (%)	Recall (%)	P	% difference
Bais et al. 2007	2,830	34.2	17.6	<0.001	16.6
Sanner et al. 2009	2,829	39.1	NA	NA	NA
Gök et al. 2010	28,073	27.5	16.6	<0.001	10.9
Gök et al. 2012	26,409	30.8	6.5	<0.001	24.3
Virtanen et al. 2011a	4,160	29.8	26.2	0.02	3.6
Virtanen et al. 2011b	8,699	31.5	25.9	<0.001	5.6
Szarewski et al. 2011	3,000	10.2	4.5	<0.001	5.7
Giorgi et al. 2011 mail-out	2,480	19.6	14.9	0.03	4.7
Giorgi et al. 2011 on demand	2,480	8.7	14.9	0.0006	–6.2
Wikstrom et al. 2011	4,060	39.0	9.0	<0.001	30.0
Tamalet et al. 2013	9,334	25.1	7.3	NR	17.8

Abbreviations: NA = not applicable, NR = not reported

Excluding this comparison between self-collection for HPV testing on demand and repeat recall, all trials reported a statistically significantly higher participation rate for self-collection. The difference in participation rates between study arms ranged from 3.6% to 30.0% in favour of self-collection (Table 107).

The Snijders et al. (2013) review identified evidence that differences in response rates for self-collection for HPV testing have been related to ethnicity and screening history. They cited evidence from Gök et al. (2012), a pooled analysis of data from two HPVself-collection studies in the Netherlands (N=15,274) that examined the impact of ethnic status, screening history and age on participation rates. The investigators reported a higher participation rate for native Dutch speakers than immigrants (34% versus 25%, P<0.001); and for women who had attended screening in the previous round versus underscreened or never screened women (34% versus 25%, P<0.001). They also found that among women aged 39 years or older who were underscreened or never screened, those that were never-screened had a higher participation rate than those that were under screened (25% versus 22%, P<0.001). They did not find an association between age and response rate (Gök et al. 2012). They also reported that the yield of CIN2+/3+ was higher for self-collection participants than screening attenders (CIN3+ relative risk 1.8 95% CI 1.5–2.1), and this association increased with age (test of homogeneity of CIN3+ yield by age, p=0.03).

Three trials reported on the proportion of non- and under-attenders who completed follow-up of abnormal results identified by self-collection for HPV testing (Szarewski et al. 2007; Gök et al. 2010; Tamalet et al. 2013).

Gök et al. (2010) reported 90.4% of 757 self-collection for HPV testing responders who had a positive HPV test attended recommended general practitioner follow-up for cytology. For 182 women with abnormal cytology at follow-up, 82.4% attended direct referral for colposcopy. For the 502 women with normal cytology, the 57% adherence to follow-up testing at one year was substantially lower than for screening attenders (86%).

Szarewski et al. (2011) reported 100% of eight self-collection for HPV testing responders who returned a positive HPV test attended recommended follow-up for cytology. Of these one woman returned a negative test and the remaining seven all took up the offer for immediate colposcopy (stage IB cancer=1, CIN2/3=2, <CIN2 or no biopsy=4).

Tamalet et al. (2013) reported 43 of 62 self-responders who returned a positive HPV test attended follow-up referral within 12 months (69.4%, 95% CI 57.0–79.4). Of these, 14 had colposcopy, 13 had a biopsy and five had surgery and overall, 36 were considered to have a 'normal cervix' (29 by cytology, 1 by colposcopy and 6 by histology).

Overall, these results provide strong evidence that a strategy of self-collection for HPV testing is feasible and effective for targeting women who would not otherwise attend screening, although participation varies in different populations.

Discussion

None of the included studies considered in the report for any of the review questions were conducted in populations vaccinated against HPV. Thus, the prevalence and incidence of cervical cancer and precancerous cervical lesions in the included studies will differ to that expected in Australia in the future, particularly in young women. Early Australian data confirm that the prevalence of the HPV genotypes vaccinated against (6, 11, 16 and 18) has decreased in vaccinated populations (Tabrizi et al. 2012), and this effect is expected to continue. Exploration of the likely effects of this on the relative effectiveness and cost-effectiveness of alternative screening approaches has been explored in an accompanying modelled analysis.

LBC-based screening has a similar accuracy to that of screening with conventional cytology, with an advantage of lower unsatisfactory rates. In previous MSAC reviews, the use of LBC in the Australian setting has been considered not cost-effective in the Australian setting. Whether this is maintained in a screening setting according to IARC recommendations is explored in a modelled analysis.

The body of evidence available to compare the relative accuracy of automated image analysis to manual LBC is poor and inconsistent. Therefore, no conclusions on its relative value can be made.

The evidence for the accuracy of liquid based cytology (LBC) is limited to comparative accuracy data and test performance measures. Accuracy conclusions are based upon a meta-analysis predominantly consisting of studies of cell filtration LBC (six of nine included studies; Arbyn et al. 2008), plus two additional studies of cell filtration LBC included in this update (Siebers et al. 2009; Klug et al. 2013). Comparisons of automated image analysis of LBC slides to manual reading of conventional cytology slides are based entirely on studies of cell filtration LBC. The body of evidence for a comparison of automated image analysis and manual reading of LBC slides is poor and inconsistent and comprises studies of both cell filtration and cell enrichment.

The evidence base for the effectiveness of HPV based screening strategies (with or without triage or co-testing) is good consisting of several large RCTs whose findings are relatively consistent for the outcomes reported despite differences in design.

Many of these trials demonstrate an increased detection of high grade CIN in the first round, followed by a decrease in detection in the second round with one RCT also demonstrating a significant decrease in the detection of cervical cancer in the second screening round, although the number of cases is small (POBASCAM). Interpretation is complex in the absence of a decrease in the cumulative detection rate of cervical cancer or its surrogate CIN3+. Increased detection in round one may represent earlier diagnosis (of uncertain benefit) and/or detection of additional disease, of which the proportion destined to progress is uncertain. Strategies in the second round differed across trials limiting the ability to interpret second round results.

Many of the HPV based screening strategies show an increase in the colposcopy rates in the first screening round. The colposcopy rates may decrease in later rounds; however these data are generally not available. Reporting of full data from these trials, including referrals to colposcopy for each round and overall, and longer follow up to allow for more complete ascertainment, may reduce the level of uncertainty regarding the trade-offs between the benefits and harms of these screening strategies.

These uncertainties regarding the lack of studies in vaccinated populations and the use of surrogate and accuracy outcomes in the existing literature are addressed by modelling long-term

benefits and harms, using an established model of cervical cancer screening, diagnosis and treatment in HPV vaccinated and unvaccinated populations in the Australian setting.

Conclusions

Safety

Conventional cytology, LBC with manual or automated slide reading and HPV testing are all safe procedures.

Any screening strategy has trade-offs between harms and benefits and the safety of screening strategies should be explored in a modelled analysis of cervical cancer screening, diagnosis and treatment under IARC recommendations and consider the impact of HPV vaccination and the potential long-term benefits and trade-offs of various screening strategies in the Australian setting.

Effectiveness

Cytology-based screening

- Cervical cancer is very rare before the age of 25 years.
- HPV vaccination is anticipated to substantially reduce the risk of cervical cancer in young women.
- There are limited comparative data on the age at which to start and stop screening.
- The available data do not demonstrate a benefit of screening women aged less than 25 years. However, screening in this age group does result in increased investigations, treatments and potential harms, without decreasing mortality.
- The majority of cervical cancer cases in women aged 65 and over are in women not meeting the criteria for an exit test.
- Extending the screening interval from two to three years is unlikely to substantially alter cervical cancer incidence or mortality rates but will lead to a reduction in the number of cytology tests and colposcopy procedures (based on two modelling studies).
- For women aged over 50 years screening intervals of five years offer risk reductions approaching those observed at three years in younger women (based on an analysis of cancer in women aged 55 to 69 years).

LBC-based screening

- LBC provides no statistically significant difference in sensitivity (at HSIL, LSIL or pLSIL thresholds) or specificity (at HSIL or LSIL thresholds) when compared with conventional cytology (MSAC 2009).
- LBC reduces the rate of unsatisfactory smears in comparison with conventional cytology (MSAC 2009).
- There is limited evidence regarding the effectiveness of automated image analysis when compared with manual reading of LBC.
- Automated image analysis with the ThinPrep Imager system detects as many CIN2+ lesions as conventional cytology, and may detect more (MSAC 2009).
- There is no evidence of an advantage, disadvantage or equivalence of the accuracy of the FocalPoint system compared to conventional cytology (MSAC 2009).
- Automated image analysis with the ThinPrep Imager system yields lower unsatisfactory rates than manual slide reading of conventional slides.

- The HPV triage test is more sensitive than a single repeat cytology test for the detection of CIN2+ and CIN3+ lesions in women with pLSIL, and has similar specificity.
- The HPV triage test is more sensitive than a single repeat cytology test for the detection of CIN2+ lesions (but not CIN3+) in women with LSIL, and has lower specificity.
- A significant proportion of additional CIN2+ lesions that would be detected by HPV triage of pLSIL and LSIL are likely to regress when a strategy of repeat cytology is used.
- The colposcopy rate following HPV triage is higher in women aged<35 years compared to women aged ≥35 years.

HPV-based screening

- HPV testing has a high negative predictive value (the probability that a negative test result is a true negative).
- The balance of benefits from increased detection of precancerous cervical lesions and potential harms from increased colposcopy referrals is more favourable for women aged ≥30–35 years for all HPV based screening strategies.
- HPV testing alone (without triaging) for primary screening is more sensitive and less specific than cytology screening for CIN2+ and CIN3+.
- HPV testing alone for primary screening increases colposcopy referrals in comparison to cytology-based screening.
- The high rate of detection of CIN2+ and CIN3+ in unvaccinated women younger than age 35 suggests that an excess of regressive lesions is identified with HPV screening in this age group (NTCC HPV testing alone trial).
- Co-testing (adjunctive or dual HPV and cytology testing) for either HPV or cytology positivity is marginally (but significantly) more sensitive than HPV testing alone.
- Co-testing for either HPV or cytology positivity is significantly less specific than HPV testing alone.
- HPV and cytology co-testing does not demonstrate a clear advantage over HPV testing alone (based on indirect comparisons which are prone to bias). Cytology triage of HPV primary testing has a higher colposcopy and retesting rate than cytology screening alone (at a referral threshold of LSIL+) in unvaccinated women <35 years.
- Cytology triage of HPV primary testing has a similar colposcopy and retesting rate to cytology screening alone (at a referral threshold of LSIL+) in women ≥35 years.
- HPV types 16 and 18 are associated with a higher risk of development of high-grade CIN than other oncogenic HPV types.
- HPV self-collection has a moderate-high sensitivity and comparably high specificity for detecting CIN2+ compared to clinic HPV testing.
- The accuracy of HPV self-collection varies for different types of sampling devices and HPV tests.
- Providing HPV self-collection kits to women who do not attend for cervical screening or who are under-attenders improves screening participation. The magnitude of this is uncertain.

Appendix A MSAC terms of reference and membership

MSAC's terms of reference are to:

- advise the Minister for Health and Ageing on the strength of evidence pertaining to new and emerging medical technologies and procedures in relation to their safety, effectiveness and cost-effectiveness and under what circumstances public funding should be supported;
- advise the Minister for Health and Ageing on which new medical technologies and procedures should be funded on an interim basis to allow data to be assembled to determine their safety, effectiveness and cost-effectiveness;
- advise the Minister for Health and Ageing on references related either to new and/or existing medical technologies and procedures; and
- undertake health technology assessment work referred by the Australian Health Ministers' Advisory Council (AHMAC) and report its findings to AHMAC.

The membership of MSAC at the November 2013 meeting comprised a mix of clinical expertise covering pathology, nuclear medicine, surgery, specialist medicine and general practice, plus clinical epidemiology and clinical trials, health economics, consumers, and health administration and planning:

Member	**Expertise**
MSAC Executive	-
Professor Robyn Ward (Chair)	Medical Oncology
Dr Frederick Khafagi (Deputy Chair)	Nuclear Medicine and Endocrinology
Professor Jim Butler (Chair Evaluation Sub-Committee)	Health Economics
Professor Andrew Wilson (Chair Protocol Advisory Sub-Committee)	Health Medicine and Epidemiology
MSAC Members	-
Associate Professor John Atherton	Cardiology
Associate Professor Michael Bilous	Anatomical Pathology
Ms Janette Donovan	Consumer Health Forum Representative
Associate Professor Kirsty Douglas	General Practice/Research
Professor Kwun Fong	Thoracic and Sleep Medicine
Professor Paul Glasziou	Evidence based Health care
Mr Scott Jansson	Medical scientist (pathology)
Professor David Little	Orthopaedics
Mr Russell McGowan	Healthcare Consumer
Associate Professor Bev Rowbotham	Haematology
Professor Richard Slaughter	Radiologist
Dr Graeme Suthers	Genetics/Pathology
Dr Simon Towler	AHMAC nominee (ex officio member), Western Australia Chief Medical Officer, part-time Intensivist
Dr Christine Tippett	Obstetrics/Gynaecology
Associate Professor David Winlaw	Paediatric Cardiothoracic Surgery

Appendix B National Cervical Cancer Screening Renewal Steering Committee

Steering committee—Application name No. 1276

Name of member	Role on committee
Professor Ian Hammond	Chair and Gynaecological Oncology expert
Professor David Roder	Epidemiology expert
Associate Professor Marion Saville	Cytology and Clinical Epidemiology expert
Professor Gordon Wright	Pathology expert
Dr Fiona Douglas	General Practice expert
Ms Lisa Peberdy	Nursing expert
Ms Tess Whittakers	Consumer advocate
Ms Jennifer Muller	Screening Programs expert
Ms Anna Burnham	NCSP Program Manager representative
Ms Louise Galloway	Standing Committee on Screening representative
Ms Alice Creelman	Department of Health and Ageing representative

Evaluators

Name	Organisation
Suzanne Dyer	NHMRC Clinical Trials Centre
Samara Lewis	NHMRC Clinical Trials Centre
Sally Lord	NHMRC Clinical Trials Centre
Anna Stoklosa	NHMRC Clinical Trials Centre
Sally Wortley	NHMRC Clinical Trials Centre

Appendix C Summary of screening programs

Country	Estimated Incidence /10^5 #	Estimated mortality /10^5 #	Population Age range	Frequency of testing in asymptomatic women: conventional cytology	Frequency of testing in asymptomatic women: LBC	Frequency of testing in asymptomatic women: HPV	HPV vaccine recommendation
Australia[a]	4.9	1.4	20–69 years *Start:* 18–20 years, or 1–2 years after first having sexual intercourse, whichever is later *Exit*: 70 years if two consecutive negative test results in the last five years	Every 2 years	-	-	Females aged 12–13[b]
Canada[e]	6.6	1.9	25–69 years *Exit:* If 3 consecutive normal test results in the last 10 years	Every 3 years	-	-	Primarily recommended in females 9–13 years but also in females 14–26 years[f]
New Zealand[m]	5.5	1.6	20–70 years Exit: If 2 consecutive negative tests	Every 3 years	-	HPV triage if ASC-US or LSIL cytology and age >30 years	Females aged 12 to 18[n]
United Kingdom[g]	7.2	2.0	25–49 years	–	Every 3 years	HPV triage if abnormal cytology	Females 12–13 years[h]
-	-	-	50–64 years Exit: If 3 consecutive normal test results, and last test after age 50 years	–	Every 5 years	HPV triage if abnormal cytology	-
United States of America[c]	5.7	1.7	21–29 years *Start*: 21 years or 3 years after first having sexual intercourse, whichever is later	Every 3 years	-	-	Females 11–12 years[d]
-	-	-	30–65 years *Exit*: If 3 consecutive negative cytology results or 2 consecutive negative HPV results in last 10 years, and most recent test in last 5 years	Every 3 years OR every 5 years with HPV co-testing	-	*Alternative*: cytology/HPV co-testing every 5 years	-

Europe

Country	Estimated Incidence /10^5 [#]	Estimated mortality /10^5 [#]	Population Age range	Frequency of testing in asymptomatic women: conventional cytology	Frequency of testing in asymptomatic women: LBC	Frequency of testing in asymptomatic women: HPV	HPV vaccine recommendation
Austria[i]	5.7	2.3	≥18 years	Every 1 year*	-	-	19–15 years of age[j]
Belgium[i]	8.4	2.7	25–64 years	Every 3 years*	-	-	10–13 years of age[j]
Bulgaria[i]	21.9	6.5	31–65 years	Every 2 years*	-	-	None[j]
Czech Republic[i]	14.0	3.9	25–69 years	Every 1 year*	-	-	None[j]
Denmark[i,l]	12.1	2.5	23–59 years	Every 3 years	-	-	12 years of age[j]
-	-	-	>45 or 50 (some counties only)	Every 5 years	-	-	-
Estonia[i,k]	15.9	6.2	30–59 years	Every 5 years, after one negative smear	--	-	None[j]
Finland[i,l]	4.5	1.2	30–60 years	Every 5 years	-	-	None[j]
France[i]	7.1	1.8	25–65 years	Every 3 years*	-	-	14 years of age[j]
Germany[i]	6.9	2.3	≥20	Every 1 year*	-	-	12–17 years of age[j]
Greece[i]	4.1	1.6	≥20	Every 1 year*	-	-	12–15 years of age[j]
Hungary[i,l]	16.6	5.6	25–65 years	Every 3 years, after one negative smear	-	-	None[j]
Iceland[l]	8.4	0.8	20–69 years	Every 2 years	-	-	12 years of age[j]
Ireland[i]	10.9	3.3	25–60 years	Every 3 years*	-	-	12 years of age[j]
-	-	-	-	Every 5 years*	-	-	-
Italy[i]	6.7	1.5	25–64 years	Every 3 years*	-	-	12 years of age[j]
Latvia[i]	12.4	7.3	20–70 years	Every 3 years*	-	-	None[j]
Lithuania[i,l]	21.0	8.3	30–60 years	Every 3 years (also reported as every 5 years[l])	-	-	None[j]
Luxembourg[i,l]	6.3	1.9	≥15 years	Every 1 year	-	-	11–12 years of age[j]
Netherlands[i,l]	6.8	1.9	30–60 years	Every 5 years	-	-	12 years of age[j]
Norway[l]	9.4	2.3	25–69 years	Every 3 years	-	-	12 years of age[j]
Poland[i]	11.6	5.8	25–59 years	Every 3 years*	-	-	None[j]
Portugal[i]	12.2	3.6	25–64 years	Every 3 years*	-	-	13 years of age[j]
Romania[i]	23.9	11.8	25–65 years	Every 5 years*	-	-	10–11 years of age[j]
Slovak Republic[i]	15.8	4.8	≥18 years	Every 1 year*	-	-	None[j]
Slovenia[i,l]	11.1	2.8	20–64 years	Every 3 years	-	-	None[j]
Spain[i,l]	6.3	1.9	20–34 years	Every 3 years (initially	-	-	11–14 years of age[j]

Country	Estimated Incidence /10[5] #	Estimated mortality /10[5] #	Population Age range	Frequency of testing in asymptomatic women: conventional cytology	Frequency of testing in asymptomatic women: LBC	Frequency of testing in asymptomatic women: HPV	HPV vaccine recommendation
				two smears 1 year apart)			
-	-	-	20–34 years	Every 3 years (initially two smears 1 year apart)	-	-	-
Sweden[i],[l]	7.8	1.9	23–50 years	Every 3 years	-	-	11–12 years of age[j]
-	-	-	51–60 years	Every 5 years	-	-	-

Abbreviations: HPV = human papillomavirus; LBC= liquid-based cytology

Age-standardised incidence rates per 100,000 GLOBOCAN 2008 (IARC). Available at http://globocan.iarc.fr/factsheets/cancers/cervix.asp#INCIDENCE. Last accessed 25 July 2013.

*Details of test method not provided.

[a] Australian Government, Department of Health and Ageing. National Cervical Screening Program. Available at http://www.health.gov.au/internet/screening/publishing.nsf/content/ncsp-policies. Last accessed 25 July 2013.

[b] Australian Government, Department of Health and Ageing. National Immunization Program Schedule. Available at http://www.immunise.health.gov.au/internet/immunise/publishing.nsf/Content/nips-ctn. Last accessed 25 July 2013.

[c] Moyer, V. Screening for cervical cancer: US Preventive Services Task Force recommendation statement. Annals of Internal Medicine. 2012;156:880–891.

[d] Saslow D, Castle PE, Cox JT et al. American Cancer Society guideline for human papillomavirus (HPV) vaccine use to prevent cervical cancer and its precursors. CA Cancer J Clin. 2007;57:7–28.

[e] Canadian Task Force on Preventive Health Care. Recommendations of Cervical Cancer. Canadian Medical Association Journal. 2013;185(1):35–45.

[f] An Advisory Committee Statement (ACS), National Advisory Committee on Immunization (NACI). Update on human papillomavirus (HPV) vaccines. Canada Communicable Disease Report. 2012;38.

[g] NHS Cervical Screening Programme. Available at http://www.cancerscreening.nhs.uk/cervical/about-cervical-screening.html. Last accessed 25 July 2013.

[h] National Health Service. Vaccinations. Available at http://www.nhs.uk/Planners/vaccinations/Pages/Teenagershub.aspx. Last accessed 25 July 2013.

[i] European Commission. Cancer screening in the European Union: Report on the implementation of the Council Recommendation on cancer screening. 2008.

[j] European Cervical Cancer Association. HPV vaccination across Europe. 2009.

[k] Veerus P, Arbyn M, Amati C, Paolo B. Impact of implementing a nationwide cervical cancer screening program on female population coverage by Pap-tests in Estonia. Tumori. 2010; 96(4): 524–8.

[l] Anttila A, Ronco G, Clifford G, Bray F, Hakama M, Arbyn M et al. Cervical cancer screening programmes and policies in 18 European countries. Br J Cancer. 2004; 91(5): 935–41.

[m] New Zealand Government. New Zealand Screening Unit. National Cervical Screening Programme. Available at: http://www.nsu.govt.nz/current-nsu-programmes/national-cervical-screening-programme.aspx. Last accessed 25 July 2013

[n] New Zealand HPV Project. Available at http://www.hpv.org.nz/patient/hpv_vaccine.htm. Last accessed 25 July 2013.

Appendix D FIGO staging of cervical cancer

The Federation Internationale de Gynecologie et d'Obstetrique (FIGO) classification system for staging cervical cancer:

The Federation Internationale de Gynecologie et d'Obstetrique (FIGO) classification system for staging cervical cancer:
Stage I
Stage I is carcinoma strictly confined to the cervix; extension to the uterine corpus should be disregarded. The diagnosis of both Stages IA1 and IA2 should be based on microscopic examination of removed tissue, preferably a cone, which must include the entire lesion.
Stage IA: Invasive cancer identified only microscopically. Invasion is limited to measured stromal invasion with a maximum depth of 5 mm and no wider than 7 mm.
Stage IA1: Measured invasion of the stroma no greater than 3 mm in depth and no wider than 7 mm diameter.
Stage IA2: Measured invasion of stroma greater than 3 mm but no greater than 5 mm in depth and no wider than 7 mm in diameter.
Stage IB: Clinical lesions confined to the cervix or preclinical lesions greater than Stage IA. All gross lesions even with superficial invasion are Stage IB cancers.
Stage IB1: Clinical lesions no greater than 4 cm in size.
Stage IB2: Clinical lesions greater than 4 cm in size.
Stage II
Stage II is carcinoma that extends beyond the cervix but does not extend into the pelvic wall. The carcinoma involves the vagina, but not as far as the lower third.
Stage IIA: No obvious parametrial involvement. Involvement of up to the upper two thirds of the vagina.
Stage IIB: Obvious parametrial involvement, but not into the pelvic sidewall.
Stage III
Stage III is carcinoma that has extended into the pelvic sidewall. On rectal examination, there is no cancer-free space between the tumour and the pelvic sidewall. The tumour involves the lower third of the vagina. All cases with hydronephrosis or a non-functioning kidney are Stage III cancers.
Stage IIIA: No extension into the pelvic sidewall but involvement of the lower third of the vagina.
Stage IIIB: Extension into the pelvic sidewall or hydronephrosis or non-functioning kidney.
Stage IV
Stage IV is carcinoma that has extended beyond the true pelvis or has clinically involved the mucosa of the bladder and/or rectum.
Stage IVA: Spread of the tumour into adjacent pelvic organs.
Stage IVB: Spread to distant organs.

Source: Sobin, L & Wittekind Ch (eds) (2002), TNM: Classification of malignant tumours, 6th edn, UICC International Union against Cancer, Geneva, Switzerland, New York: Wiley-Liss, pp. 155–157.

Appendix E Cervical cancer incidence and mortality in Australia (1983–2007)

Table 108 Data for Figure 3: Overall incidence and mortality from cervical cancer in Australia

Time period	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
Incidence	14.0	12.8	9.4	7.2	6.8

Time period	1983–1987	1988–1992	1993–1997	1998–2002	2003–2007
Mortality	4.8	4.2	3.5	2.4	2.0

Note: Mortality and incidence per 100,000 Australian women

Table 109: Data for Figure 4: Age-stratified incidence from cervical cancer in Australia, 1985–2009

Age group	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
15–19	0.4	0.1	0.2	0.2	0.1
20–24	2.7	1.8	1.5	1.6	1.4
25–29	10.1	7.5	7.0	5.8	6.7
30–34	19.6	16.0	11.6	9.9	10.5
35–39	21.2	20.2	15.4	10.8	11.4
40–44	22.1	21.1	15.7	10.1	11.3
45–49	22.7	20.8	14.6	12.8	10.5
50–54	20.4	19.5	13.1	10.4	10.1
55–59	20.4	20.4	13.6	9.6	8.7
60–64	25.4	21.6	16.5	11.4	9.6
65–69	28.1	25.1	17.5	12.3	11.4
70–74	25.0	24.6	16.9	12.3	10.0
75–79	24.0	23.0	18.0	14.2	11.3
80–84	24.5	23.2	20.0	17.8	13.6

Note: incidence per 100,000 Australian women

Table 110: Data for Figure 5: Age-stratified mortality of cervical cancer in Australia, 1985–2009

Age group	1983–1987	1988–1992	1993–1997	1998–2002	2003–2007
15–19	0.0	0.1	0.0	0.0	0.0
20–24	0.2	0.1	0.0	0.2	0.1
25–29	1.1	0.8	0.4	0.4	0.5
30–34	2.3	2.1	1.4	1.0	1.1
35–39	3.1	3.3	2.5	1.6	1.3
40–44	4.4	4.4	3.5	2.3	2.2
45–49	6.1	6.3	4.9	3.2	2.3
50–54	7.2	4.8	5.3	3.6	2.9
55–59	9.4	7.1	6.5	3.4	2.7
60–64	10.7	9.4	6.9	5.7	4.8
65–69	16.3	11.6	9.5	5.9	4.4
70–74	15.4	14.4	11.9	8.5	4.9
75–79	14.7	14.8	13.8	9.1	7.4
80–84	20.7	17.7	15.0	12.6	9.7

Note: Mortality per 100,000 Australian women

Appendix F Search strategies

Research question 1: Screening age range and interval

EMBASE.com search (EMBASE and Medline databases):

1. 'uterine cervix tumor'/exp
2. 'uterine cervix dysplasia'/exp
3. #1 OR #2
4. 'mass screening'/exp
5. 'uterine cervix cytology'/exp
6. screen*:ab,ti
7. #4 OR #5 OR #6
8. #3 AND #7
9. 'cervical cancer screening':ab,ti
10. 'cervical screening':ab,ti
11. 'cervical neoplasm screening':ab,ti
12. 'cervical neoplasia screening':ab,ti
13. #11 OR #12
14. #8 OR #9 OR #10 OR #13
15. 'coitus'/exp
16. (first NEAR/4 intercourse):ab,ti
17. (first NEAR/4 coitus):ab,ti
18. (initi* NEAR/4 intercourse):ab,ti
19. (sexual* NEAR/4 activ*):ab,ti
20. 'chronologic age':ab,ti
21. 'chronological age':ab,ti
22. #20 OR #21
23. 'different age':ab,ti
24. 'different ages':ab,ti
25. #23 OR #24
26. (young* NEAR/2 wom*n):ab,ti
27. (age NEAR/2 specific):ab,ti
28. (beg*n* NEAR/4 screen*):ab,ti
29. (start* NEAR/4 screen*):ab,ti
30. (age NEAR/4 beg*n*):ab,ti
31. (age NEAR/4 start*):ab,ti
32. (age NEAR/4 first):ab,ti
33. 'age factors'/exp
34. 'age distribution'/exp
35. (old* NEAR/2 wom*n):ab,ti
36. (stop* NEAR/4 screen*):ab,ti
37. (age NEAR/4 stop*):ab,ti
38. 'age restrict$':ab,ti
39. (withdraw* NEAR/4 screen*):ab,ti
40. #15 OR #16 OR #17 OR #18 OR #19 OR #22 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40

41. #14 AND #41
42. [english]/lim
43. [humans]/lim
44. [2010-2013]/py
45. #42 AND #43 AND #44 AND #45

PubMed search:

1. uterine cervical diseases/ or uterine cervical dysplasia/ or uterine cervical neoplasms/
2. Cervical Intraepithelial Neoplasia/
3. Vaginal Smears/
4. 1 or 2 or 3
5. mass screening/
6. screen$.ti,ab.
7. 5 or 6
8. 4 and 7
9. cervical cancer screening.ti,ab.
10. cervical neoplas$ screening.ti,ab.
11. cervical screening.ti,ab.
12. 8 or 9 or 10 or 11
13. Coitus/
14. (first adj4 intercourse).ti,ab.
15. (first adj4 coitus).ti,ab.
16. (initi$ adj4 intercourse).ti,ab.
17. (sexual$ adj4 activ$).ti,ab.
18. chronologic$ age.ti,ab.
19. different age$.ti,ab.
20. (young$ adj2 wom#n).ti,ab.
21. (age adj2 specific).ti,ab.
22. (beg#n$ adj4 screen$).ti,ab.
23. (start$ adj4 screen$).ti,ab.
24. (age adj4 beg#n$).ti,ab.
25. (age adj4 start$).ti,ab.
26. (age adj4 first).ti,ab.
27. age factors/
28. age distribution/
29. (old$ adj2 wom#n).ti,ab.
30. (stop$ adj4 screen$).ti,ab.
31. (age adj4 stop$).ti,ab.
32. age restrict$.ti,ab.
33. (withdraw$ adj4 screen$).ti,ab.
34. 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33
35. 12 and 34
36. limit 35 to english language

limit 36 to yr="2010 - 2013"

Research question 2: Liquid-based cytology

EMBASE.com search (EMBASE and Medline databases):

1. ‘cancer screening’/syn
2. screening:de, ab,ti
3. 1 OR 2
4. ‘uterine cervix cancer’/syn OR ‘uterine cervix tumor’/syn
5. 4 AND 3
6. ‘uterine cervix cytology’/syn OR ‘papanicolaou test’/syn
7. (liquid OR cervical OR cervix) NEAR/3 cytolog*):ab,ti
8. 5 OR 6 OR 7
9. ‘thin prep’ OR ‘thin layer’ or thin*prep:dn,ab,ti
10. liquid:ab,ti OR ‘fluid’:ab,ti
11. Cytorich:ab,dn,ti OR autocyte:ab,dn,ti OR ‘sure path’ OR sure*path:ab,dn,ti
12. 9 OR 10 OR 11
13. 8 AND 12
14. 14 AND (2010-2012)/py
15. ‘automation’/syn
16. comput* OR automat*
17. ‘image processing’/syn OR ‘image cytometry’/syn
18. focal*point:dn,ab,ti OR ‘focal point’ OR ‘papnet’ or ‘cytosavant’ OR ‘autopap’:dn,ab,ti
19. 16 OR 17 OR 18
20. 8 and 19
21. 20 and (2008-2012)/py

PubMed

1. Cervical intraepithelial neoplasia [MeSH Terms]
2. ‘uterine cervical neoplasms [MeSH Terms]
3. Screening [Title/Abstract]
4. Mass screening [MeSH Terms]
5. Vaginal smears [MeSH Terms]
6. liquid based [Title]
7. Thinprep [Title/Abstract]
8. (Computer or Automated [Title]) and Cervical [Title]
9. focal point [Title/Abstract]
10. surepath [Title/Abstract]
11. 1 or 2
12. 3 or 4 or 4
13. 6 or 7 or 10
14. 12 and 13
15. 11 and 13
16. 14 or 15 publication 2010 – 2012
17. Diagnosis, Computer-Assisted [MAJR]
18. Diagnosis, Computer-Assisted/instrumentation [MAJR]
19. Cytodiagnosis/instrumentation [MAJR]
20. Cytological Techniques/instrumentation [MAJR]
21. 17 or 18 or 19 and 20 or 9
22. 19 and 11
23. 21 and 13
24. 22 or 23 or 8 publication 2008-2012

Research question 3: HPV testing

EMBASE.com search (EMBASE and Medline databases):

25. 'papanicolaou test'/syn OR 'papanicolaou test'
26. 'cancer screening'/syn OR 'cancer screening' OR screening:ab,ti AND ('uterine cervix cancer'/syn OR 'uterine cervix cancer' OR 'uterine cervix tumor'/syn OR 'uterine cervix tumor')
27. 'uterine cervix cytology'/syn OR 'uterine cervix cytology'
28. #1 OR #2 OR #3
29. 'papilloma virus'/syn OR 'papilloma virus' OR 'DNA probe'/syn OR 'DNA probe'
30. digene*:ab,ti OR 'hybrid capture':ab,ti OR HC*:ab,ti
31. triage*:ab,ti
32. (HPV* NEAR/3 test*):ab,ti
33. (HPV* NEAR/3 detect*):ab,ti
34. #5 OR #6 OR #7 OR #8 OR #9
35. $4 AND #10
36. #11 AND [1-9-2010]/sd

PubMed search

1. Search ("vaginal smears"[MeSH Terms] OR "early detection of cancer"[MeSH Terms] OR "uterine cervical neoplasms"[MeSH Terms])
2. Search "papilloma virus"[title/abstract] OR "digene"[title/abstract] OR "hybrid capture"[title/abstract] OR "HC"[title/abstract] OR "triage"[MeSH Terms] OR "papillomavirus infections/diagnosis"[MAJR]
3. Search (#1) AND (#2)
4. Search HPV[Title]
5. Search "mass screening"[MAJR]
6. Search (#4) AND (#5)
7. Search (#3) OR (#6)

   Search (#3) OR (#6) Filters: Publication date from 2010/09/01 to 2012/12/04

Appendix G Existing HTA reports and systematic reviews 2008–2012

Table 111 Relevant existing HTAs and systematic reviews published since the 2009 MSAC reviews (2009–present)

Organisation Country (if applicable)	Date	Title	Scope
Pantanwala et al	2013	A systematic review of randomised trials assessing human papillomavirus testing in cervical cancer screening	HPV screening strategies in RCTS Search to: 2010 Relevant to present review : Question 3 Identified post-search, no additional studies beyond AHRQ
Racey et al	2013	Self-collected HPV Testing Improves Participation in Cervical Cancer Screening: A Systematic Review and Meta-analysis	Attendance of self-collection Search to: July 2012 Relevant to present review: Question 3 Identified post-search, doesn't alter findings of Snijders 2012.
Arbyn et al Cochrane Review	2013	Human Papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions (Review)	HPV as triage Search to: January 2011 Relevant to present review: Question 2 Included
Snijders et al	2013	High-risk HPV testing on self-sampled versus clinician-collected specimens: a review on the clinical accuracy and impact on population attendance in cervical cancer screening	Accuracy, attendance of self-collection Search to: January 2011 Relevant to present review's: Question 3 Included
Chen et al	2012	Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis.	Visual inspection, HPV DNA testing, manual LBC testing. Search to: August 2011 Relevant to present review's: Question 2, Question 3 Search dates superseded by Arbyn 2012
ACS-ASCCP-ASCP Cervical Cancer Guidelines Committee (Saslow et al) United States	2012	American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer.	HPV DNA as primary test, HPV DNA and cytology as co-tests, HPV with reflex cytology, cytology with HPV DNA triage. Search to: July 2011 Relevant to present review's: Question 2, Question 3 Included
Correa et al	2012	Colposcopic triage methods for detecting cervical intraepithelial neoplasia grade 3 after cytopathological diagnosis of low-grade squamous intraepithelial lesion: a systematic review on diagnostic tests.	HPV as triage. Search to: July 2009 Relevant to present review's: Question 2 Search dates superseded by Arbyn 2013
Canadian Taskforce on Preventative Health HTA Peirson et al. Canada	2012	Screening for cervical cancer	Effect of cervical screening on incidence and mortality of invasive cervical cancer (incl. LBC-based methods, HPV DNA testing, varying screening interval, entry and exit from the screening programme), harms of cervical screening, effect of screening in sub-groups, patient values and preferences, self-collection. Search to: April 2012 (key questions), February 2011 (contextual questions) Relevant to present review's: Question 1, Question 2, Question 3 Included

Organisation Country (if applicable)	Date	Title	Scope
Malaysian Ministry of Health - Health Technology Assessment Section (MaHTAS) Malaysia	2012	HPV DNA-based screening test for cervical cancer	Assessment of HPV test as primary test, in triage, in clinical management, and as follow-up to detect treatment failure. Search to: August 2012 Relevant to present review's: Question 3 Search dates superseded by AHRQ / US PSTF
Arbyn et al 2012	2012	Evidence Regarding Human Papillomavirus Testing in Secondary Prevention of Cervical Cancer	HPV & cytology co-testing, HPV vs. cytology testing. Search to: December 2011 (triage), February 2012 (HPV primary) Relevant to present review's: Question 3 Included
Hayes, Inc. United States	2012	Human papillomavirus (HPV) genotyping to test for the presence of HPV 16 and 18 in women	Sensitivity and specificity of HPV genotyping for HPV 16, 18, early stage cervical cancer. Patient management with HPV genotyping. Search dates unknown Relevant to present review's: Question 3 Report not freely available
American College of Gynaecologists (ACOG) United States	2012	Screening for cervical cancer	Screening start, frequency, and exit (age & testing). HPV testing, cytology (conventional and LBC) testing, and co-testing. Search to: March 2012 Relevant to present review's: Question 1, Question 2, Question 3 Report not freely available NB: these guidelines supersede the ACOG 2006/2009 guidelines
Fontaine et al	2012	Unsatisfactory rates vary between cervical cytology samples prepared using ThinPrep and SurePath platforms: a review and meta-analysis	Compares unsatisfactory rates between two LBC platforms (ThinPrep and SurePath) Search to: August 2011 Relevant to present review's: Question 2
Murphy et al Canada	2012	HPV Testing in Primary Cervical Screening: A Systematic Review and Meta-Analysis	HPV testing vs. cytology testing in primary screening. Search to: November 2010 Relevant to present review's: Question 3 Narrower scope than AHRQ
Ronco et al Italy	2012	Use of liquid-based cytology for cervical cancer precursors screening	Compares manually-read CC and manually-read LBC Search to: NR Relevant to present review's: Question 2 In Italian
Dalla Palma et al Italy	2012	Computer-assisted Pap test for cervical cancer screening	Compares manually-read CC and automatically-read CC. Search to: NR Relevant to present review's: Question 2, Question 3 In Italian
Ronco et al	2012	Health technology assessment report: HPV DNA based primary screening for cervical cancer precursors	HPV primary screening Search to: NR Relevant to present review: Question 3 Results in English, remainder in Italian.
IQWiG Germany	2011	Benefit assessment of HPV test in primary screening for cervical cancer.	Scope unknown (German only)Search to: July 2011 Relevant to present review's: Question 3 In German
AHRQ / US PSTF (Vesco et al) United States	2011	Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force	Age of screening initiation; manual LBC compared to CC; HPV DNA testing for primary screening without LBC triage; HPV DNA testing for primary screening with LBC triage; harms of LBC; harms of HPV DNA testing. Search to: September 2010 Ancillary search to: August 2011 Relevant to present review's: Question 1, Question 2,

Organisation Country (if applicable)	Date	Title	Scope
			Question 3 Included
University of Michigan Health System United States	2011	Cancer Screening	Screening initiation, frequency, and exit. Cytology (conventional and LBC) vs. HPV; HPV& cytology co-testing. Search to: June 2010 Relevant to present review's: Question 1, Question 2, Question 3. Superseded
DIMDI (Sroczynski et al) Germany	2010	Decision-analytic modelling to evaluate the long-term effectiveness and cost-effectiveness of HPV-DNA testing in primary cervical cancer screening in Germany.	Long-term clinical and economic consequences of HPV DNA testing in primary cervical screening. Search to: n/a Relevant to present review's: Question 3 In German (only a summary available in English)
Zhao et al China	2010	Performance of high-risk human papillomavirus DNA testing as primary screen for cervical cancer: a pooled analysis of individual patient data from 17 population studies from China.	HPV as primary screening test (compared to cytology and VIA). Search to: n/a; studies between 1999-2008 in China Relevant to present review's: Question 3 Wrong population
Institute of Health Economics (Ospina et al) Canada	2009	Human papillomavirus (HPV) testing in Alberta.	HPV as primary test (alone or co-testing with CC or LBC); HPV as triage test for women with borderline abnormalities (for colposcopy and biopsy referral). Search to: March 2008 Relevant to present review's: Question 3 Superseded

Table 112 Relevant on-going studies identified through clinical trials databases and the Renewal Steering Commitee

Principal investigator	Location	Population Age	Approx. size (target)	Investigation	Planned completion	Primary outcomes
Lytwyn et al	Canada	35-69 yrs	1440	Self-collection vs. 2nd reminder letter	Jan 2013	Uptake of screening invitation
Kulasingam et al	United States	21 yrs+	5400	Cytology vs. self-sample (HPV)	Aug 2016	Sensitivity and specificity of self-collected at home human papillomavirus test with reflex cytology. Total estimated costs associated with cervical cancer screening. Patient satisfaction with self-collected at home human papillomavirus test screening.
Ngan et al	Hong Kong	30-60 yrs	12,000	CC vs. HPV + CC co-testing	June 2017	Histological CIN2, CIN3 and cervical carcinoma
Dillner et al	Sweden	56-65 yrs	Not stated	CC vs. HPV	Not stated	number of CIN 2+ cases for the two different diagnostic procedures
Canfell and Saville (Compass trial)	Australia	25-69 yrs	Not stated (pilot of 5000)	CC with HPV triage vs. HPV with genotyping triage	Not stated	CIN3+
Simpson et al (iPAP trial)	Australia	Not stated	Not stated	home-based HPV self-collection vs. reminder letter	2015	Participation and follow up

Appendix H Excluded studies

LBC

Incorrect outcomes

Beerman H, van Dorst EB, Kuenen-Boumeester V, Hogendoorn PC. Superior performance of liquid-based versus conventional cytology in a population-based cervical cancer screening program. *GynecolOncol.* 2009 Mar; 112(3):572-6 [reference standard threshold for positive CIN1+]

Fregnani JH, Scapulatempo C, Haikel RL, Mauad EC, Campacci N, Longatto-Filho A. Liquid-based cytology improves detection of cervical intraepithelial lesion in Low and High-risk women for HPV related diseases. Abstract presented at the Global Academic Program (GAP), 14-16 May 2012, Oslo, Norway [no accuracy data]

Longatto-Filho A, Naud P, Derchain SFM, Roteli-Martins C, Tatti S, Hammes LS, Sarian LO, Erzen M, Branca M, De Matos JC, Gontijo R, Maeda MYS, Lima T, Costa S, Syrjanen S, & Syrjanen K (2012). Performance characteristics of Pap test, VIA, VILI, HR-HPV testing, cervicography, and colposcopy in diagnosis of significant cervical pathology. *Virchows Arch* 460, 577-585. [no accuracy data; non-randomised]

Maccallini V, Angeloni C, Caraceni D, Fortunato C, Venditti MA, Gabriele G, Antonelli C, Lattanzi A, Puliti D, Ciatto S, Confortini M, Sani C, Zappa M. Comparison of the conventional cervical smear and liquid-based cytology: Results of a controlled, prospective study in the Abruzzo Region of Italy. *Acta Cytologica*, 2008; 52(5):568-574 [accuracy only reported for ASCUS+ threshold]

Wright PK, Marshall J, & Desai M (2010). Comparison of SurePath(R) and ThinPrep(R) liquid-based cervical cytology using positive predictive value, atypical predictive value and total predictive value as performance indicators. *Cytopathology* 21, 374-378. [relative sensitivity not valid as not randomised; cohort study; superseded]

Zhao FH, Hu SY, Bian JJ, Liu B, Peck RB, Bao YP, Pan QJ, Frappart L, Sellors J, & Qiao YL (2011). Comparison of ThinPrep and SurePath liquid-based cytology and subsequent human papillomavirus DNA testing in China. *Cancer Cytopathol* 119, 387-394. [absolute accuracy measures not valid as no verification of negative test results; insufficient raw data to calculate relative accuracy measures]

Study design lower level of evidence than included studies

Esquivias Lopez-Cuervo J, Montalban Beltran E, Cuadros Lopez JL, Alonso Castillo A, & Nieto Sanchez T (2011). Preliminary study of a new, fully automated system for liquid-based cytology: The NovaPrep(registered trademark) processor system. *Acta Cytol* 55, 281-286. [split-sample; LBC not commercially available (NovaPrep)]

Hariri J (2011). Evaluation of liquid-based cytology in public screening for cervical cancer in Southern Jutland in Denmark. *APMIS* 119, 44-45. [historical comparison]

Suwannarurk K, Bhamarapravatana K, Thaweekul Y, Mairaing K, Poomtavorn Y, & Pattaraarchachai J (2011). A 1-year experience with liquid-based and conventional papanicolaou

smear in Thammasat University Hospital. *J Med Assoc Thai* 94 Suppl 7, S47-S51. (Issue 12) [retrospective, non-randomised study]

Taoka H, Yamamoto Y, Sakurai N, Fukuda M, Asakawa Y, Kurasaki A, Oharaseki T, & Kubushiro K (2010). Comparison of conventional and liquid-based cytology, and human papillomavirus testing using SurePath preparation in Japan. *Hum Cell* 23, 126-133. [split-sample; incomplete verification of positive results]

Other

Castle PE, Bulten J, Confortini M, Klinkhamer P, Pellegrini A, Siebers AG, Ronco G, & Arbyn M (2010). Age-specific patterns of unsatisfactory results for conventional Pap smears and liquid-based cytology: data from two randomised clinical trials. BJOG : an international journal of obstetrics and gynaecology 117, 1067-1073. [unsatisfactory rate data from NTCC and NETHCON – additional analyses of included data]

Strander B, Andersson-Ellström A, Milsom I, Rådberg T, Ryd W. Liquid-based cytology versus conventional Papanicolaou smear in an organized screening program : a prospective randomized study. Cancer, 2007; 111(5):285-291 [psuedorandomised trial with poor randomisation CC:LBC allocation 1.9 : 1 designed as 1:1, adjusted OR for detection used as outcome, but not for HSIL threshold]

Automated image analysis

Incorrect outcomes

Bowditch RC, Clarke JM, Baird PJ, & Greenberg ML (2011). Results of an Australian trial using SurePath liquid-based cervical cytology with Focalpoint computer-assisted screening technology. *Diagn Cytopathol*. [no accuracy outcomes]

Levi AW, Chhieng DC, Schofield K, Kowalski D, & Harigopal M (2012). Implementation of FocalPoint GS location-guided imaging system: experience in a clinical setting. *Cancer Cytopathol* 120, 126-133. [no accuracy outcomes; historical comparison]

Pacheco MC, Conley RC, Pennington DW, & Bishop JW (2008). Concordance between original screening and final diagnosis using imager vs. manual screen of cervical liquid-based cytology slides. *Acta Cytol* 52, 575-578.

Quddus M, Neves T, Reilly M, Steinhoff M, & Sung C (2009). Does the ThinPrep Imaging System increase the detection of high-risk HPV-positive ASC-US and AGUS the Women and Infants Hospital experience with over 200,000 cervical cytology cases. *CytoJournal* 6.

Renshaw AA & Elsheikh TM (2011). Predicting screening sensitivity from workload in gynecologic cytology: A review. *Diagn Cytopathol* 39, 832-836.

Renshaw AA & Elsheikh TM (2012). HSIL, epithelial cell abnormality-adjusted workload, and the Thinprep imaging system. *Diagn Cytopathol* **40**, 201-203.

Renshaw AA & Elsheikh TM (2012). Low grade squamous intraepithelial lesion, epithelial cell abnormality-adjusted workload, and the ThinPrep imaging system. *Diagn Cytopathol* **40**, 698-700.

Thrall MJ, Russell DK, Bonfiglio TA, & Hoda RS (2008). Use of the ThinPrep(registered trademark) Imaging System does not alter the frequency of interpreting Papanicolaou tests as atypical squamous cells of undetermined significance. *CytoJournal* 5.

Wrong use of technology

Heard T, Chandra A, Culora G, Gupta SS, Herbert A, & Morgan M (2012). Use of the ThinPrep Imaging System for internal quality control of cervical cytology. *Cytopathology*. [QC use]

Wilgenbusch H, Mueller G, Neal M, & Renshaw AA (2011). Rapid prescreening is as effective at reducing screening error as postscreening with the FocalPoint automated screening device. *Diagn Cytopathol* **39**, 818-821.

Not in English

Soler Font I, Romero Martos E, Pijuan Andujar L, Lloveras Rubio B, Carreras Collado R, Serrano Figueras S, & Alameda Quitllet F (2010). The use of automated readings in gynaecological cytology. The cytotechnician's viewpoint. *Revista Esp Patol* 43, 69-72.

Study design lower level of evidence than included studies

Papillo JL, St.John TL, & Leiman G (2008). Effectiveness of the ThinPrep Imaging System: Clinical experience in a low risk screening population. Diagn Cytopathol 36, 155-160. [historical comparison]

HPV

Incorrect outcomes

Schiffman M, Glass AG, Wentzensen N, Rush BB, Castle PE, Scott DR et al. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. *Cancer Epidemiol Biomarkers Prev* 2011; 20(7):1398-1409. [For genotyping – no accuracy data]

Wright J, Stoler MH, Behrens CM, Apple R, Derion T, Wright TL. The ATHENA human papillomavirus study: Design, methods, and baseline results. *Am J Obstet Gynecol* 2012; 206(1):46. [For genotyping – no accuracy data]

Study design lower level evidence than included studies

Cuzick J, Thomas Cox J, Zhang G, Einstein MH, Stoler M, Trupin S et al. Human papillomavirus testing for triage of women with low-grade squamous intraepithelial lesions. *Int J Cancer* 2012. [cohort study]

Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: A population-based study in routine clinical practice. *Lancet Oncol* 2011; 12(7):663-672. [cohort study]

Kjaer SK, Frederiksen K, Munk C, Iftner T. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *J Natl Cancer Inst* 2010. [cohort study]

Levi AW, Harigopal M, Hui P, Schofield K, Chhieng DC. Use of high-risk human papillomavirus testing in patients with low-grade squamous intraepithelial lesions. *Cancer Cytopathol* 2011; 119(4):228-234. [cohort study]

Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Rozendaal L, Heideman DA et al. HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. *Br J Cancer* 2012; 106(5):975-981. [cohort study]

Soderlund-Strand A, Eklund C, Kemetli L, Grillner L, Tornberg S, Dillner J et al. Genotyping of human papillomavirus in triaging of low-grade cervical cytology. *Am J Obstet Gynecol* 2011; 205(2):145-146. [cohort study, wrong population for genotyping]

Stoler MH, Wright TC, Jr., Sharma A, Apple R, Gutekunst K, Wright TL. High-risk human papillomavirus testing in women with ASC-US cytology: results from the ATHENA HPV study. *Am J Clin Pathol* 2011. [cohort study]

Stoler MH, Wright TC, Jr., Sharma A, Zhang G, Apple R, Wright TL et al. The interplay of age stratification and HPV testing on the predictive value of ASC-US cytology. Results from the ATHENA HPV study. *Am J Clin Pathol* 2012; 137(2):295-303. [cohort study]

Abstract only

Cotton S, Sharp L, Little J, Cruickshank M, Seth R, Smart L et al. The role of human papillomavirus testing in the management of women with low-grade abnormalities: Multicentre randomized controlled trial. *Obstet Gynecol Surv* 2010; 65(7):432-433.

Appendix I Cell enrichment liquid-based cytology

In Australia two LBC platforms/systems are available for cervical cancer screening: cell filtration (ThinPrep) LBC and cell enrichment (SurePath) LBC. These systems use different technical methods for storing and preparing the cervical cytology sample, some of which are patented. The SurePath™ LBC system (Beckton Dickinson Pty Ltd) requires that the head of the brush or spatula to be detached into a vial of liquid to produce a cell suspension which undergoes "enrichment" prior to slide preparation via gravity sedimentation. The ThinPrep® Pap system (Hologic [Australia] Pty Ltd) requires that the head of the brush or spatula to be rinsed into a vial of liquid to produce a cell suspension which then undergoes a membrane "filtration" and the cell residue is transferred to the slide.

Both the cell filtration and cell enrichment Pap systems can be reviewed using either manual or automated reading methods.

The 2009 MSAC Assessment Report (MSAC 1122 AR) considered both these methods in its assessment of LBC. While literature was included on both cell filtration and cell enrichment, the majority of the included literature, and those studies that were considered more applicable, reported on cell filtration. This review provides an update of the literature published since 2009 on LBC (both cell enrichment and cell filtration), as well as an update of the literature for automated image analysis.

No studies reporting on manual LBC cell enrichment compared to conventional cytology met the inclusion criteria for this review. As such the evidence for manual LBC is based entirely on cell filtration evidence. Two included studies providing a comparison of automated image analysis using the FocalPoint system (based on cell enrichment slides) to manual LBC were included (MAVARIC, Kitchener et al. 2011a; Wilbur et al. 2009).

One excluded study by Bowditch (2011) compared BD FocalPoint GS computer-assisted screening of cell enrichment BD SurePath (registered trademark) liquid-based cervical cytology slides (SP + FP) to conventional cytology in an Australian setting. It was a split sample trial of 2,198 routine specimens with seeded cases and thus was a study providing a lower level of evidence for manual LBC than the included studies. In addition, accuracy outcomes were not reported. The rate of unsatisfactory specimens in the automation assisted arm was 0.2% compared with 4.1% in the conventional Pap test, a significant reduction. Appendix H lists the reasons that studies were excluded from the review. This list includes a number of studies that report on cell enrichment LBC (Beerman et al. 2009; Zhao et al. 2011; Taoka et al. 2010). Broadly speaking the studies that reported on cell enrichments were excluded from this review because of study design (emphasis was placed on identifying the highest quality evidence on LBC rather than a specific system), or the study did not include relevant outcome data for this review (incidence and mortality of cervical cancer, relative or absolute accuracy of CIN2/CIN3).

While this review does not include evidence on manual cell enrichment LBC; cell enrichment LBC (specifically the BD SurePath™ LBC system) was the subject of a submission based assessment that underwent assessment by MSAC in April 2013. MSAC considered evidence on cell enrichment LBC in respect to cervical cancer screening. MSAC accepted that, as found by MSAC in 2009 for application 1122, the technique is as safe as conventional cytology and there is no basis to conclude a significant clinical difference between cell enrichment LBC and cell filtration LBC

and conventional cytology. Public funding was not supported for the proposed specific listing of cell enrichment LBC, which excluded cell filtration LBC.

Unsatisfactory rates have not been systematically addressed in the ARHQ report (Vesco et al. 2011) but were considered in the 2009 MSAC report. Unsatisfactory rates were also listed in the DAP for the assessment of cell enrichment LBC, with the evidence report for this procedure considered by MSAC in April 2013.

As part of the literature search for this review, a study was identified that sought to compare the unsatisfactory rates between the two major liquid-based cytology (LBC) platforms (Fontaine et al. 2012). It has not been included in the main body of this review but has been included here for completeness. The authors of this study conducted a systematic review and meta-analysis of the literature, searching articles in PubMed published in English between January 1990 and August 2011. Forty-two studies (14 cell enrichment and 28 cell filtration) were included in the quantitative analysis and four studies in the meta-analysis. The four studies included in the meta-analysis presented data in the same population by the same laboratory for both cell enrichment and cell filtration LBC methodologies.

It was reported that the pooled unsatisfactory rate of the 1,120,418 cervical cytology smears in the 14 different studies using cell enrichment LBC, was 0.3% (non-comparative data). Using cell filtration LBC, 1,148,755 smears reported from 28 studies determined a pooled unsatisfactory rate of 1.3% (Fontaine et al. 2012). When controlling for other variables there was no difference between the systems; however, the power to detect a difference was low. The meta-analysis demonstrated cell enrichment LBC to have a significantly lower unsatisfactory rate compared with cell filtration LBC with a pooled relative risk of 0.44 (95% CI 0.25–0.77).

No sensitivity analysis was undertaken despite considerable variation being noted in the unsatisfactory rates among jurisdictions. The authors also note that the systematic review and meta-analysis only looks at one aspect of LBC system selection and other factors such as accuracy must also be taken into consideration.

Appendix J Previous MSAC assessments

Ref 12a (2002) – Liquid-based cytology for cervical screening

A submission was made to the MSAC to consider the merits of replacing the conventional cytology with liquid-based cytology (LBC) as the cervical screening test for possible listing on the Medicare Benefits Schedule (MBS). Cervical screening utilising LBC was intended to replace conventional cytology for detecting cancerous lesions and also precancerous cells so that treatment may be initiated before the disease progresses to an inoperable stage.

Aim

The aim of the assessment was to assess the safety, effectiveness and cost-effectiveness of LBC for cervical screening.

Method

MSAC conducted a systematic review of medical literature using the Cochrane Library, Medline, PreMedline, Current Contents, Biological Abstracts, CINAHL and EMBASE databases from January 2000 to April 2002 to identify the accuracy and precision of the tests and their usefulness in terms of patient outcomes in the context of the current Australian cervical screening guidelines. Assessment of clinical effectiveness relied on five secondary studies and seven primary studies. Assessment of cost-effectiveness was based on both review of a submitted model and revision of this model on the basis that LBC was no worse than conventional cytology in detecting high-grade lesions.

Conclusions and results

Safety: No risks were associated with the test itself, although the safety issues were the same as those for conventional cytology because the method of collecting cellular material is the same for both.

Effectiveness: There was insufficient evidence to draw definitive conclusions regarding the diagnostic characteristics of LBC and conventional cytology for cervical screening. The lack of high-quality evidence on the performance of LBC did not permit evaluation of whether it was equal or superior in effectiveness to conventional cytology. The assessment concluded that further high-quality studies using an acceptable reference standard, such as histological confirmation of cytology results, would be crucial in allowing a valid and reliable judgment concerning the sensitivity and specificity of LBC.

Cost-effectiveness: A decision analytic model indicated that LBC was associated with greater costs per woman than conventional cytology. Since there was insufficient evidence to support a claim that LBC is superior to conventional cytology in detecting high-grade lesions or invasive cancer, it followed that there is no evidence to suggest that LBC would be cost-effective at the proposed price.

Recommendations

MSAC advised that there was insufficient evidence to support public funding of liquid-based cytology for cervical screening at the time of the assessment.

Application 1122 (2009)—Automation-Assisted and Liquid-Based Cytology for Cervical Cancer Screening

Aim

The primary research question was:

- What is the safety, effectiveness and cost-effectiveness of liquid-based cytology (LBC) using automated image analysis systems in comparison to manual reading of conventionally prepared Pap smear cytology samples for the screening and diagnosis of cervical cancer?

The following secondary research questions were also addressed in terms of the safety, effectiveness and cost-effectiveness of:

- LBC compared to conventionally prepared Pap smear cytology samples when manual reading of slides is used?
- automated image analysis systems in comparison to manual reading of conventionally prepared Pap smear cytology samples?
- LBC using automated image analysis systems compared to manual reading of LBC?

Methods

A systematic review was undertaken of the medical literature published since the last MSAC report in 2002 – up to 6 February 2008 – on the evidence for automation-assisted and liquid-based cytology for cervical cancer screening.

Results and conclusions

Safety: The assessment reported noted that LBC with manual or automation-assisted slide reading uses the same procedure for collecting cervical cell samples as conventional Pap cytology tests and is considered a safe procedure.

Effectiveness: No studies assessed the impact of LBC with manual or automated slide reading on the incidence of invasive cervical cancer or consequent mortality rates compared to conventional cytology. The review therefore relies on evidence about the accuracy of manual or automated LBC for detecting precancerous cervical lesions to draw conclusions about its relative effectiveness. The report concluded that manual LBC compared to conventional cytology provides no statistically significant increase in sensitivity or specificity; provides no statistically significant difference in sensitivity (HSIL, LSIL or pLSIL thresholds) or specificity (HSIL or LSIL thresholds) for the detection of CIN2+; and reduces the specificity for the detection of CIN2+ at a test threshold of pLSIL. Automated LBC detects more CIN2+ lesions compared to conventional cytology, but results from one trial raises uncertainty about whether this difference is attributable to LBC alone, to the automation-assisted reading system or a combination of both. Both manual and automated LBC classify more slides as positive for low-grade lesions and both reduce the rate of unsatisfactory smears. Automated LBC also reduces slide processing time.

Economic considerations: A modelled analysis based on favourable assumptions regarding test characteristics found that automated LBC (ThinPrep Imager) would be associated with a cost of $194,835 per life year saved (LYS). Manual LBC was associated with a cost of $126,315 per LYS to $385,982 per LYS depending on the level of reimbursement. The analysis predicted automated LBC would prevent 68 cervical cancer cases and 19 deaths due to cervical cancer annually, and that manual LBC would prevent 23 cervical cancer cases and six deaths due to cervical cancer annually. Both would also result in additional follow-up procedures. Each cancer case averted with automated LBC would require an additional 566 cytology tests, 159 colposcopies, 76 biopsies and 26 treatments for CIN2/3. Each cancer case averted with manual LBC would require an additional 990 cytology tests, 295 colposcopies, 142 biopsies and 32 treatments for CIN2/3. The findings are sensitive to assumed relative test accuracy; differences in the unsatisfactory smear rate; assumptions about disease natural history (particularly for high-grade regression and progression); the recommended screening interval; and the cost of the new technology. The results presented were based on the current screening program in Australia without taking into account potential changes resulting from HPV vaccination.

Recommendations

MSAC's advice was, with respect to LBC, that in comparison to conventional cytology, LBC is safe, is at least as effective, but is not cost-effective at the price requested and advised that LBC should not be supported for public funding.

With respect to automated (computerised) testing of LBC, that in comparison to conventional cytology, automated LBC testing is safe, is at least as effective, but is not cost-effective at the price requested. MSAC advised that automated testing of LBC specimens should not be supported for public funding.

Ref 12b (2002) – Human Papillomavirus Testing in Women with Cytological Prediction of Low-grade Abnormality

Aim

To assess the safety, effectiveness and cost-effectiveness of HPV testing in women with cytological prediction of low-grade abnormality.

The report presented an evaluation of the clinical effectiveness and cost-effectiveness of the HC-II HPV DNA (HC-II) test in women with cytological prediction of low-grade abnormality. It integrated submissions made to the MSAC from the technology manufacturer and a request made to the MSAC from within the Commonwealth Department of Health and Ageing.

A submission was made to the MSAC to consider the merits of adjunctive HPV testing as an aid for diagnosing the presence of high-risk HPV subtypes in women with screen-detected low-grade abnormality for possible listing on the Medicare Benefits Schedule (MBS).

Method

MSAC conducted a systematic review of medical literature using the Cochrane Library, Medline, PreMedline, Current Contents, Biological Abstracts, CINAHL and EMBASE databases from January 1999–April 2002 to identify the accuracy of the test and its usefulness in the context of the current Australian cervical screening guidelines. Assessment of clinical effectiveness relied on one secondary study and seven primary studies. Assessment of cost-effectiveness was based on both a

review of a submitted model and a decision analytic model simulating the cost-effectiveness of HPV testing in women with cytological prediction of low-grade abnormality.

Conclusions and results

Safety: No risks were associated with the test itself. Safety issues were the same as those for conventional cytology because the method of collecting cellular material is the same.

Effectiveness: A systematic review concluded that HPV testing was more sensitive but less specific than cytology, although current evidence would not support widespread implementation. Although it was inappropriate to calculate an overall summary statistic as primary studies reported varying positive thresholds for HPV testing, pooling four of seven studies revealed a sensitivity of 92% (95% CI 87–97%) and specificity of 54% (95% CI 50–57%) in detecting lesions that were moderately dysplastic or worse. It was noted that these results should be interpreted with caution as studies failed to meet all validity criteria, which may result in non-appraisable bias. The assessment concluded that additional high-quality studies using an acceptable reference standard, such as histological confirmation of cytology results, would be useful in allowing a valid and reliable judgment of the sensitivity and specificity of HPV testing in this population.

Cost-effectiveness: A decision analytic model indicated that HPV testing was both more expensive and less effective in detecting high-grade lesions than the management plan currently recommended by the NHMRC, but the model was particularly sensitive to the estimated prevalence of high-grade lesions in women.

Recommendations

MSAC advice was that there was currently insufficient evidence to support public funding at the time for the use of the HPV test for triaging of women with equivocal cervical screening results.

Ref 12d (2003) – Human Papillomavirus Testing for Cervical Cancer

Aim

The aim of the assessment was to assess the safety, effectiveness and cost-effectiveness of human papillomavirus (HPV) testing by the Hybrid Capture 2 (HC2) test for cervical screening as either a stand-alone screening test or combined with screening by cytology. The report reviewed the detection of high-risk HPV subtypes by the HC2 test for routine cervical screening either as a stand-alone screening test or as an adjunct to either conventional cytology or liquid-based smear.

Method

MSAC conducted a systematic review of medical literature using the Cochrane Library, Medline, PreMedline, Current Contents, Biological Abstracts, CINAHL and EMBASE databases from January 1998 to October 2002 to identify the accuracy and precision of the tests and their usefulness in terms of patient outcomes in the context of the current Australian cervical screening guidelines. Assessment of clinical effectiveness relied on two primary studies, while assessment of cost-effectiveness was based on review of a submitted economic model.

Conclusions and results

Safety: No risks were associated with the test itself, although the safety issues were the same as those for conventional cytology because the method of collecting cellular material is the same for both.

Effectiveness: There was insufficient evidence that HPV testing is effective in detecting high-grade cervical lesions when used as either a stand-alone screening test or combined with screening by cytology.

Cost-effectiveness: Due to insufficient evidence of clinical effectiveness, an economic evaluation could not be performed.

Recommendations

MSAC advised that there was insufficient evidence to support public funding of HPV testing as a stand-alone screening test or as an adjunct to cervical cytology screening.

Further research

The assessment noted that at the time, three trials were underway, enrolling more than 43,000 women with results expected over the next three years: HART (UK), ARTISTIC (UK) and CCaST (Canada).

Ref 39 (2009) – Human Papillomavirus Triage Test for Women with Possible or Definite Low-Grade Squamous Intraepithelial Lesions

Aim

The aim of this assessment was to assess the safety, effectiveness and cost-effectiveness of the Hybrid Capture 2 (HC2) human papillomavirus (HPV) triage test for asymptomatic women with a routine screening cervical cytology test finding of possible or definite low-grade squamous intra-epithelial lesions (LSIL) when testing is conducted either at the time of the index cytology result, or with repeat cytology 12 months after the index cytology test, compared with the standard strategy of repeating cervical cytology annually for two years. In addition, assessment of the HPV triage test was undertaken for subgroups of women, with an index cytology finding of definite LSIL (dLSIL) only; possible LSIL (pLSIL) only; aged ≥23 years; and aged ≥30 years.

Methods

A systematic review was undertaken of the medical literature published since the last MSAC report in 2002 – up to 6 February 2008 – to report on the evidence addressing the research questions. This evidence was used to identify 18 potential HPV triage strategies for further assessment. A comprehensive model of cervical screening, diagnosis and treatment was developed to estimate and compare the incremental effectiveness and cost-effectiveness of adopting each of these strategies in Australia.

Results and conclusions

Safety: The assessment concluded that the HC2 HPV test is a technically safe procedure.

Effectiveness: No published clinical studies comparing the impact of the HPV triage test on invasive cervical cancer incidence or mortality rates compared with the current standard strategy of repeat cytology testing of women with an index cytology finding of pLSIL or dLSIL were identified. Comparative accuracy studies provided strong evidence that an immediate HC-II HPV triage test is a more sensitive test than a single repeat cytology test for detecting cervical intraepithelial neoplasia (CIN) 2+ lesions in women with pLSIL, and has similar specificity. There was also strong evidence that an immediate HC2 HPV triage test is no more sensitive than a single repeat cytology test for detecting CIN2+ lesions in women with dLSIL, and has lower specificity, but colposcopy referral rates may be favourable compared with a strategy of two annual repeat cytology tests in this patient group. Restricting the HPV triage test to older age groups is associated with a higher specificity and lower colposcopy referral rate and a corresponding smaller gain in sensitivity compared with its use in all age groups. No published clinical studies comparing the accuracy of performing the HC2 HPV test with repeat cytology at the 12-month follow-up visit with the current strategy were identified.

Economic considerations: The modelled analysis predicted that, compared with current practice, a strategy of performing the HPV triage test for women aged 30+ years produces an incremental cost-effectiveness ratio (ICER) of $75,739 per life year saved (LYS) if conventional cytology is used with co-collection for HPV testing; or $83,496 per LYS using manual liquid-based cytology (LBC) (at an incremental price of $2.40 compared with conventional cytology) with reflex HPV testing; or $170,209 per LYS using automated LBC with reflex HPV testing. When compared with current practice using conventional cytology, the most cost-effective options were strategies involving immediate triage either with conventional cytology and co-collection or with reflex manual LBC testing (at an incremental price of $2.40 compared with conventional cytology). The evaluation found that performing immediate HPV triage is substantially more cost-effective than delaying triage testing until a 12-month follow-up visit, for all types of cytology (conventional with co-collection, manual LBC and automated LBC). The ICER estimates were most sensitive to assumptions concerning the cost of the HPV test and cytology tests including the cost of co-collection, the HPV test characteristics, the discount rate, the recommended screening interval, and the likelihood that CIN3 lesions will progress to cancer. The results presented were based on the current screening program in Australia without taking into account potential changes resulting from HPV vaccination.

Recommendation

MSAC's advice was that HPV triage testing in cervical cancer was not cost-effective in the Australian setting at the current price of HPV testing and did not support public funding.

Ref 12c (2003) – Computer-assisted Image Analysis for Cervical Screening

Aim

To assess the safety, effectiveness and cost-effectiveness of computer-assisted image analysis for cervical screening cytology compared with manual processing. The comparator for computer-assisted image analysis was conventional manual screening in which trained laboratory personnel examine the slides using light microscopy.

Method

MSAC conducted a systematic review of medical literature using the Cochrane Library, Medline, PreMedline, Current Contents, Biological Abstracts, CINAHL and EMBASE databases from January 1966 to September 2002 to identify the accuracy and precision of the tests and their usefulness in terms of patient outcomes. This report adopted the criteria for assessment of validity of evidence recommended by the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests.

Conclusions and results

Safety: Computer-assisted image analysis is conducted in the laboratory on the same slides as conventional cervical cytology. The safety issues with this technology were the same as those for the cytological assessment methods in current use.

Effectiveness: There was insufficient evidence to assess whether computer-assisted image analysis is as effective as manual processing for cervical screening cytology.

Cost-effectiveness: Due to insufficient evidence of clinical effectiveness, an economic evaluation could not be performed.

Recommendations

MSAC advised that there was insufficient evidence to support public funding of computer-assisted image analysis for cervical screening at this time.

Appendix K Public consultation

The Draft Review of Evidence was released for public consultation on 7 June 2013 and closed on 4 July 2013. The report was posted on Department of Health and Ageing's website and the Partner Reference Group and other key stakeholders were notified about the consultation process via an email letter. Feedback on the Draft Review of Evidence was provided by the following individuals/organisations:

1. Lyn Dean – SA Health
2. Lesley Wilkinson – South Australia NCSP Program Manager
3. Gail Ward – Tasmanian SCOS Member/NCSP Program Manager
4. Lea Rawlings – Cancer Council Victoria
5. Dorota Gertig and Kelly Drennan – Victorian Cytology Service
6. Alexandra Barratt – University of Sydney
7. Carol Bennett – Consumer Health Forum of Australia
8. Bronwen Ross – The Royal College of Pathologist of Australasia
9. Annabelle Farnsworth, Jennifer Roberts, Clare Biro, Suzanne Hyne, Adele Richards - GynaePath Douglass Hanly Moir Pathology
10. Penny Stewart – Becton Dickinson
11. Stuart Dobson – Hologic (Australia)
12. Debra Brown

Glossary and abbreviations

ABS	Australian Bureau of Statistics
ACS	American Cancer Society
AGUS	atypical glandular cells of undetermined significance
AHRQ	Agency for Healthcare Research and Quality
AIHW	Australian Institute of Health and Welfare
AIS	adenocarcinoma in situ
ALTS	ASCUS–LSIL Triage Study Group
AMBS	Australian Modified Bethesda System
ARTG	Australian Register of Therapeutic Goods
ASC-H	atypical squamous cells, possible high-grade lesion
ASCCP	American Society for Colonoscopy and Cervical Pathology
ASCP	American Society for Clinical Pathology
ASCUS	atypical squamous cells, undetermined significance
CADTH	Canadian Agency for Drugs and Technologies in Health
CC	conventional cytology
CI	confidence interval
CIN	cervical intraepithelial neoplasia
CTFPHC	Canadian Task Force on Preventive Health Care
DAP	Decision Analytic Protocol
DNA	Deoxyribonucleic acid
DoHA	Department of Health and Ageing
ERSC	Evidence Review Synthesis Centre
FDA	Food and Drug Administration
FIGO	International Federation of Gynecology and Obstetrics
FP	false positive
HART	HPV in Addition to Routine Testing

HC2	Hybrid Capture 2
HPV	human papillomavirus
HSIL	high-grade squamous intraepithelial lesion
HTA	health technology assessment
IARC	International Agency for Research on Cancer
ICER	incremental cost-effectiveness ratio
ITT	intention-to-treat
IVD	in-vitro diagnostic
LBC	liquid-based cytology
LEEP	loop electrosurgical excision procedure
LLETZ	large loop excision of the transformation zone
LSIL	low-grade squamous intraepithelial lesion
LYS	life year saved
MBS	Medicare Benefits Schedule
MSAC	Medical Services Advisory Committee
NCSP	National Cervical Screening Program
NETHCON	Netherlands ThinPrep Versus Conventional Cytology
NFR	not for review
NHMRC	National Health and Medical Research Council
NHPVVP	National Human Papillomavirus Vaccination Program
NNTH	number needed to harm
NR	not reported
NTCC	New Technologies in Cervical Cancer
NSW	New South Wales
OR	odds ratio
PCR	polymerase chain reaction
pHSIL	possible high-grade squamous intraepithelial lesion

PIP	Practice Incentives Program
pLSIL	possible low-grade squamous intraepithelial lesion
POBASCAM	Population Based Screening Study Amsterdam
PPV	positive predictive value
PTR	Pap Test Registry
QC	quality control
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	randomised controlled trial
RNA	Ribonucleic acid
RR	relative risk
SCC	squamous cell carcinoma
SIL	squamous intraepithelial lesion
TGA	Therapeutic Goods Administration
TN	true negative
TOMBOLA	Trial of Management of Borderline and Other Low-grade Abnormal
TP	true positive
TPI	ThinPrep Imager
UK	United Kingdom
US	United States
USPSTF	United States Preventive Services Task Force
VCCR	Victorian Cervical Cytology Register

References

Ackerson K, Pohl J, Low L. Personal influencing factors associated with pap smear testing and cervical cancer. *Policy Polit Nurs Pract* 2008; 9(1):50-60.

Andrae B, Kemetli L, Sparen P, Silfverdal L, Strander B, Ryd W et al. Screening-preventable cervical cancer risks: evidence from a nationwide audit in Sweden. *J Natl Cancer Inst* 2008; 100(9).

Anttila A, Hakama M, Kotaniemi-Talonen L, Nieminen P. Alternative technologies in cervical cancer screening: a randomised evaluation trial. *BMC Public Health* 2006; 6(1):1-8.

Anttila A, Kotaniemi-Talonen L, Leinonen M, Hakama M, Laurila P, Tarkkanen J 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* 2010; 340.

Anttila A, Pokhrel A, Kotaniemi-Talonen L, Hakama M, Malila N, Nieminen P. Cervical cancer patterns with automation-assisted and conventional cytological screening: A randomized study. *Int J Cancer* 2011; 128(5):1204-1212.

Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second edition--summary document. *Ann Oncol* 2010; 21(3):448-458.

Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: A systematic review and meta-analysis. *Obstet Gynecol* 2008; 111(1):167-177.

Arbyn M, Martin-Hirsch P, Wentzensen N. Human papillomavirus-based triage of women showing a cervical cytology result of borderline or mild dyskaryosis. *BJOG: An International Journal of Obstetrics & Gynaecology* 2010; 117(6):641-644.

Arbyn M, Roelens J, Simoens C, Buntinx F, Paraskevaidis E, Martin-Hirsch PP et al. Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions. *Cochrane Database Syst Rev* 2013; 3:CD008054.

Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012; 30:Suppl-99.

Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE. How to evaluate emerging technologies in cervical cancer screening? *Int J Cancer* 2009; 125(11):2489-2496.

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, Supplement 3(0):S78-S89.

ASCUS-LSIL Triage Study (ALTS) Group. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. *American Journal of Obstetrics and Gynecology* 2003a; 188(6):1393-1400.

ASCUS-LSIL Triage Study (ALTS) Group. Human Papillomavirus Testing for Triage of Women With Cytologic Evidence of Low-Grade Squamous Intraepithelial Lesions: Baseline Data From a Randomized Trial. *J Natl Cancer Inst* 2000; 92(5):397-402.

ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. *American Journal of Obstetrics and Gynecology* 2003b; 188(6):1383-1392.

Australian Government, Department of Human Services. Practice Incentives Program Cervical Screening Incentive guidelines July 2012. http://www.medicareaustralia.gov.au/provider/incentives/pip/files/9519-1208en.pdf . 2012. 15-8-2013. Australian Health Ministers' Advisory Council (AHMAC). Population based screening framework. 2008. Canberra, Commonwealth of Australia.

Australian Institute of Health and Welfare (AIHW). ACIM (Australian Cancer Incidence and Mortality) Books. 2012b. Canberra, AIHW.

Australian Institute of Health and Welfare (AIHW). Cervical cancer and safety monitoring guidelines. AIHW: Canberra [2013b [cited 2013 Aug. 2]; Available from: URL: http://www.aihw.gov.au/cervical-cancer-and-safety-monitoring-guidelines/

Australian Institute of Health and Welfare (AIHW). Cervical screening in Australia 2009-2010. Cancer series no. 67. Cat. no. CAN 63. 2012a. Canberra, AIHW.

Australian Institute of Health and Welfare (AIHW). Cervical screening in Australia 2010-2011. Cancer series 76. Cat. no. CAN 72. 2013a. Canberra, AIHW.

Bais AG, van Kemenade FJ, Berkhof J, Verheijen RH, Snijders PJ, Voorhorst F et al. Human papillomavirus testing on self-sampled cervicovaginal brushes: an effective alternative to protect nonresponders in cervical screening programs. *Int J Cancer* 2007; 120(7).

Barghouti F, Takruri A, Froelicher E. Awareness and behavior about Pap smear testing in family medicine practice. *Saudi Med J* 2008; 27(9):1036-1040.

Belinson J, Du H, Yang B, Wu R, Belinson S, Qu X et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. Int J Cancer 2012; 130(8):1855-1860.

Belinson J, Hu S, Niyazi M, Pretorius R, Wang H, Wen C et al. Prevalence of type-specific human papillomavirus in endocervical, upper and lower vaginal, perineal and vaginal self-collected specimens: Implications for vaginal self-collection. Int J Cancer 2010; 127(5):1151-1157.

Belinson J, Qiao YL, Pretorius R, Zhang WH, Elson P, Li L et al. Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. *Gynecol Oncol* 2001; 83(2).

Belinson JL, Qiao YL, Pretorius RG, Zhang WH, Rong SD, Huang MN et al. Shanxi Province cervical cancer screening study II: self-sampling for high-risk human papillomavirus compared to direct sampling for human papillomavirus and liquid based cervical cytology. *Int J Gynecol Cancer* 2003; 13(6).

Bentley J. Colposcopic management of abnormal cervical cytology and histology. *Journal of Obstetrics & Gynaecology Canada: JOGC* 2012; 34(12):1188-1202.

Binns PL, Condon JR. Participation in cervical screening by Indigenous women in the Northern Territory: a longitudinal study. *Med J Aust* 2006; 185(9):490-494.

Biscotti CV, Dawson AE, Dziura B, Galup L, Darragh T, Rahemtulla A et al. Assisted primary screening using the automated ThinPrep Imaging System. *Am J Clin Pathol* 2005; 123(2):281-287.

Bjerre P, Silfverdal L, Dillner L, Hagmar Br, Edvardsson H, Dillner J et al. A randomized trial of basing treatment on human papillomavirus and/or cytology results in low-grade cervical lesion triage. *American Journal of Obstetrics and Gynecology* 2008; 199(1):24.

Bolger N, Heffron C, Regan I, Sweeney M, Kinsella S, McKeown M et al. Implementation and evaluation of a new automated interactive image analysis system. *Acta Cytol* 2006; 50(5):483-491.

Bosch FX, Lorincz A, MuÃ±oz N, Meijer CJLM, Shah KV. The causal relation between human papillomavirus and cervical cancer. *Journal of Clinical Pathology* 2002; 55(4):244-265.

Bowditch RC, Clarke JM, Baird PJ, Greenberg ML. Results of an Australian trial using SurePath liquid-based cervical cytology with Focalpoint computer-assisted screening technology. *Diagn Cytopathol* 2011.

Brotherton JML. How much cervical cancer in Australia is vaccine preventable? A meta-analysis. *Vaccine* 2008; 26(2):250-256.

Brotto L, Chou A, Singh T, Woo J. Reproductive health practices among Indian, Indo-Canadian, Canadian East Asian, and Euro-Canadian women: the role of acculturation. *J Obstet Gynaecol Can* 2008; 30(3):229-238.

Bulkmans NWJ, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJP, Bulk S et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *The Lancet* 1924; 370(9601):1764-1772.

Bulkmans NWJ, Rozendaal L, Snijders PJF, Voorhorst FJ, Boeke AJ, Zandwijken GRJ 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.

Canadian Task Force on Preventive Health Care. Recommendations on screening for cervical cancer. *Canadian Medical Association Journal* 2013; 185(1):35-45.

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. *Med J Aust* 2006; 185(9):482-486.

Castle PE, Schiffman M, Wheeler CM, Solomon D. Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2. *Obstet Gynecol* 2009; 113(1):18-25.

Castle PE, Stoler MH, Solomon D, Schiffman M. The relationship of community biopsy-diagnosed cervical intraepithelial neoplasia grade 2 to the quality control pathology-reviewed diagnoses: an ALTS report. *Am J Clin Pathol* 2007; 127(5):805-815.

Castle PE, Stoler MH, Wright TC, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical

cancer screening of women aged 25 years and older: A subanalysis of the ATHENA study. *Lancet Oncol* 2011; 12(9):880-890.

Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness: CRD's guidancce for those carrying out or commissioning reviews. 2001. York, NHS Centre for Reviews and Dissemination.

Chang AR. Carcinoma in situ of the cervix and its malignant potential. A lesson from New Zealand. *Cytopathology* 1990; 1(6).

Chang S, Woo J, Gorzalka B, Brotto L. A questionnaire study of cervical cancer screening beliefs and practices of Chinese and Caucasian mother-daughter pairs living in Canada. *J Obstet Gynaecol Can* 2010; 32(3):254-262.

Chen C, Yang Z, Li Z, Li L. Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis. Int J Gynecol Cancer 2012; 22(6):908-21.

Confortini M, Bonardi L, Bulgaresi P, Cariaggi MP, Cecchini S, Ciatto S et al. A feasibility study of the use of the AutoPap screening system as a primary screening and location-guided rescreening device. *Cancer* 2003; 99(3):129-134.

Coory MD, Fagan PS, Muller JM, Dunn NA. Participation in cervical cancer screening by women in rural and remote Aboriginal and Torres Strait Islander communities in Queensland. *Med J Aust* 2002; 177(10):544-547.

Correa F, Russomano F, Oliveira C. Colposcopic triage methods for detecting cervical intraepithelial neoplasia grade 3 after cytopathological diagnosis of low-grade squamous intraepithelial lesion: a systematic review on diagnostic tests. Sao Paulo Med J 2012; 130(1):44-52.

Cotton S, Sharp L, Little J, Cruickshank M, Seth R, Smart L et al. The role of human papillomavirus testing in the management of women with low-grade abnormalities: Multicentre randomised controlled trial. *BJOG Int J Obstet Gynaecol* 2010; 117(6):645-659.

Creighton P, Lew JB, Clements M, Smith M, Howard K, Dyer S 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.

Cuzick J, Szarewski A, Cubie H, Hulman G, Kitchener H, Luesley D et al. Management of women who test positive for high-risk types of human papillomavirus: The HART study. *Lancet* 2003; 362(9399):1871-1876.

Davey E, Barratt A, Irwig L, Chan SF, Macaskill P, Mannes P et al. Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: A systematic review. *Lancet* 2006; 367(9505):122-132.

Davey E, D'Assuncao J, Irwig L, Macaskill P, Chan SF, Richards A et al. Accuracy of reading liquid based cytology slides using the ThinPrep Imager compared with conventional cytology: Prospective study. *Brit Med J* 2007; 335(7609):31-35.

Day S, van Dort P, Tay-Teo K. Improving participation in cancer screening programs: a review of social cognitive models, factors affecting participation, and strategies to improve participation. Victorian Cytology Service: Carlton South.; 2010.

Della Palma P, Moresco L, Giorgi Rossi P. [Computer-assisted Pap test for cervical cancer screening]. Epidemiol Prev 2012; 36(5 Suppl 3):e1-e43.

Department of Health and Ageing (DoHA). National Cervical Screening Program: NCSP Policies. Australian Government [2011 [cited 2013 July 31]; Available from: URL: http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/hpv-vaccinated-policy

Dickinson J. Cervical screening: making the right change is more important than concern about discord. *Canadian Medical Association Journal* 2013.

Dickinson JA. Cervical screening: time to change the policy. *Med J Aust* 2002; 176(11):547-550.

Dillner L, Kemetli L, Elfgren K, Bogdanovic G, Andersson P, Carlsten-Thor A et al. Randomized healthservices study of human papillomavirus-based management of low-grade cytological abnormalities. *Int J Cancer* 2011; 129(1):151-159.

Dinkelspiel H, Fetterman B, Poitras N, Kinney W, Cox JT, Lorey T et al. Screening history preceding a diagnosis of cervical cancer in women age 65 and older. *Gynecol Oncol* 2012; 126(2):203-206.

Elfgren K, Rylander E, RÃ¥dberg T, Strander Br, Strand A, Paajanen K et al. Colposcopic and histopathologic evaluation of women participating in population-based screening for human papillomavirus deoxyribonucleic acid persistence. *American Journal of Obstetrics and Gynecology* 2005; 193(3):650-657.

Everett T, Bryant A, Griffin MF, Martin-Hirsch Pierre PL, Forbes CA, Jepson RG. Interventions targeted at women to encourage the uptake of cervical screening. *Cochrane Database of Systematic Reviews: Reviews* 2011; Issue 5.

Fahs MC, Mandelblatt J, Schechter C, Muller C. Cost Effectiveness of Cervical Cancer Screening for the Elderly. *Ann Intern Med* 1992; 117(6):520-527.

Fiebig D, Haas M, Hossain I, Street D, Viney R. Decisions about Pap tests: what influences women and providers? *Soc Sci Med* 2009; 68(10):1766-1774.

Foley G, Alston R, Geraci M, Brabin L, Kitchener H, Birch J. Increasing rates of cervical cancer in young women in England: An analysis of national data 1982-2006. *Brit J Cancer* 2011; 105(1):177-184.

Fontaine D, Narine N, Naugler C. Unsatisfactory rates vary between cervical cytology samples prepared using ThinPrep and SurePath platforms: a review and meta-analysis. *BMJ Open* 2012; 2(2).

Frederiksen ME, Lynge E, Rebolj M. Women's preferences regarding options for management of atypical, borderline and low-grade cervical abnormalities. *Cytopathology* 2012; 23(3):201.

Giorgi Rossi P, Marsili LM, Camilloni L, Iossa A, Lattanzi A, Sani C et al. The effect of self-sampled HPV testing on participation to cervical cancer screening in Italy: a randomised controlled trial (ISRCTN96071600). *Br J Cancer* 2011; 104(2).

Gök M, Heideman DA, van Kemenade FJ, Berkhof J, Rozendaal L, Spruyt JW 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.

Gök M, van Kemenade FJ, Heideman DA, Berkhof J, Rozendaal L, Spruyt JW et al. Experience with high-risk human papillomavirus testing on vaginal brush-based self-samples of non-attendees of the cervical screening program. *Int J Cancer* 2012; 130(5).

Gram IT, Macaluso M, Stalsberg H. Incidence of cervical intraepithelial neoplasia grade III, and cancer of the cervix uteri following a negative Pap-smear in an opportunistic screening. *Acta Obstet Gynecol Scand* 1998; 77(2).

Guyatt G, Oxman A, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336.

Halford JA, Batty T, Boost T, Duhig J, Hall J, Lee C et al. Comparison of the sensitivity of conventional cytology and the ThinPrep imaging system for 1,083 biopsy confirmed high-grade squamous lesions. *Diagn Cytopathol* 2010; 38(5):318-326.

Hartmann K, Hall S, Nanda K, Boggess J, Zolnoun D. Screening for Cervical Cancer. Systematic Evidence Reviews, No. 25. 2002. Rockville MD, Agency for Healthcare Research and Quality (US).

Hellsten C, Lindqvist P, Sjostrom K. A longitudinal study of sexual functioning in women referred for colposcopy: a 2-year follow up. *BJOG* 2008; 115(2):205-211.

Herbert A, Holdsworth G, Kubba A. Why young women should be screened for cervical cancer: the distinction between CIN2 and CIN3. *Int J Cancer* 2010; 126(9).

Hoffman M, Cooper D, Carrara H, Rosenberg L, Kelly J, Stander I et al. Limited Pap screening associated with reduced risk of cervical cancer in South Africa. *Int J Epidemiol* 2003; 32(4).

Holanda FJ, Castelo A, Veras T, de Almeida F, Lins M, Dores G. Primary screening for cervical cancer through self sampling. Int J Gynecol Obstet 2006; 95(2):179-184.

Holowaty P, Miller AB, Rohan T, To T. Natural history of dysplasia of the uterine cervix. *J Natl Cancer Inst* 1999; 91(3).

Howard K, Salkeld G, McCaffery K, Irwig L. HPV triage testing or repeat Pap smear for the management of atypical squamous cells (ASCUS) on Pap smear: is there evidence of process utility? *Health Economics* 2008; 17(5):593-605.

IARC Working Group On Evaluation Of Cervical Cancer Screening Programmes. Screening For Squamous Cervical Cancer: Duration Of Low Risk After Negative Results Of Cervical Cytology And Its Implication For Screening Policies. *British Medical Journal (Clinical Research Edition)* 1986; 293(6548):659-664.

IARC. Cervix cancer screening/IARC Working Group on the Evaluation of Cancer-Preventive Strategies. IARC Handbooks of Cancer Prevention [2005 [cited 2013 July 27]; 10 Available from: URL: http://www.iarc.fr/en/publications/pdfs-online/prev/handbook10/index.php

Ikenberg H, Harlfinger W, Neis K, Konig J, Klug S. A randomized trial comparing conventional cytology to liquid-based cytology with computer-assistance: Results of the rhine-saar-study. *Cytopathology* 2011; 22:77-78.

Inna N, Phianmongkhol Y, Charoenkwan K. Sexual function after loop electrosurgical excision procedure for cervical dysplasia. *J Sex Med* 2010; 7(3):1291-1297.

Insinga RP, Glass AG, Rush BB. Diagnoses and outcomes in cervical cancer screening: a population-based study. *Am J Obstet Gynecol* 2004; 191(1).

Irwig L, Macaskill P, Farnsworth A, Wright RG, McCool J, Barratt A et al. A randomized crossover trial of PAPNET for primary cervical screening. *Journal of Clinical Epidemiology* 2004; 57(1):75-81.

Jakobsson M, Tarkkanen J, Auvinen E, Hakkinen R, Laurila P, Tapper AM. Colposcopy referral rate can be reduced by high-risk human papillomavirus triage in the management of recurrent atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion cytology in Finland. *Int J STD AIDS* 2012; 23(7):485-489.

Ji C, Chen M, Sun J, Liang W. Cultural views, English proficiency and regular cervical cancer screening among older Chinese American women. *Womens Health Issues* 2010; 20(4):272-278.

Johnson C, Mues K, Mayne S, Kiblawi A. Cervical cancer screening among immigrants and ethnic minorities: a systematic review using the Health Belief Model. *J Low Genit Tract Dis* 2008; 12(3):232-241.

Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: A population-based study in routine clinical practice. *Lancet Oncol* 2011; 12(7):663-672.

Khan MJ, Castle PE, Lorincz AT, Wacholder S, Sherman M, Scott DR et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Natl Cancer Inst* 2005; 97(14):1072-1079.

Kitchener HC, Almonte M, Gilham C, Dowie R, Stoykova B, Sargent A et al. ARTISTIC: A randomised trial of human papillomavirus (HPV) testing in primary cervical screening. *Health Technol Assess* 2009; 13(51):1-126.

Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol* 2009; 10:672-682.

Kitchener HC, Almonte M, Wheeler P, Desai M, Gilham C, Bailey A et al. HPV testing in routine cervical screening: cross sectional data from the ARTISTIC trial. *Br J Cancer* 2006; 95(1):56-61.

Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R et al. MAVARIC - a comparison of automation-assisted and manual cervical screening: a randomised controlled trial. *Health Technol Assess* 2011c; 15(3):iii-xi, 1.

Kitchener HC, Blanks R, Dunn G, Gunn L, Desai M, Albrow R et al. Automation-assisted versus manual reading of cervical cytology (MAVARIC): A randomised controlled trial. *Lancet Oncol* 2011a; 12(1):56-64.

Kitchener HC, Gilham C, Sargent A, Bailey A, Albrow R, Roberts C 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* 2011b; 47(6):864-871.

Kjaer SK, Frederiksen K, Munk C, Iftner T. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *J Natl Cancer Inst* 2010; 102(19):1478-1488.

Klug SJ, Neis KJ, Harlfinger W, Malter A, König J, Spieth S et al. A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. *Int J Cancer* 2013; 132(12):2849-2857.

Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskevaidis E. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non-randomized studies. *Gynecol Oncol* 2007; 104(1):232-246.

Kotaniemi-Talonen L, Anttila A, Malila N, Tarkkanen J, Laurila P, Hakama M et al. Screening with a primary human papillomavirus test does not increase detection of cervical cancer and intraepithelial neoplasia 3. *Eur J Cancer* 2008; 44(4):565-571.

Kotaniemi-Talonen L, Nieminen P, Anttila A, Hakama M. Routine cervical screening with primary HPV testing and cytology triage protocol in a randomised setting. *Br J Cancer* 2005; 93(8):862-867.

Krahn M, McLachlin M, Pham B, Rosen B, Sander B, Grootendorst P et al. Liquid-based techniques for cervical cancer screening: systematic review and cost-effectiveness analysis. 2008.

Kulasingam SL, Havrilesky L, Ghebre R, Myers ER. Screening for Cervical Cancer: A Decision Analysis for the U.S. Preventive Services Task Force. AHRQ Publication No. 11-05157-EF-1. 2011. Rockville, MD, Agency for Healthcare Research and Quality.

Kulasingam SL, Hughes JP, Kiviat NB, Mao C, Weiss NS, Kuypers JM et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: Comparison of sensitivity, specificity, and frequency of referral. *J Am Med Assoc* 2002; 288(14):1749-1757.

Kyrgiou M, Koliopoulos G, Martin-Hirsch P, Arbyn M, Prendiville W, Paraskevaidis E. Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: systematic review and meta-analysis. *Lancet* 2006; 379(9509):489-498.

Kyrgiou M, Koliopoulos G, Martin-Hirsch P, Kehoe S, Flannelly G, Mitrou S et al. Management of minor cervical cytological abnormalities: A systematic review and a meta-analysis of the literature. *Cancer Treat Rev* 2007; 33(6):514-520.

Kyrgiou M, Stasinou SM, Arbyn M, Valasoulis G, Ghaem MS, Martin-Hirsch-Pierre PL et al. Management of low-grade squamous intra-epithelial lesions of the uterine cervix: repeat cytology versus immediate referral to colposcopy. *Cochrane Database Syst Rev* 2012.

Lancucki L, Sasieni P, Patnick J, Day TJ, Vessey MP. The impact of Jade Goody's diagnosis and death on the NHS Cervical Screening Programme. *J Med Screen* 2012; 19(2):89-93.

Leinonen M, Nieminen P, Kotaniemi-Talonen L, Malila N, Tarkkanen J, Laurila P et al. Age-Specific Evaluation of Primary Human Papillomavirus Screening vs Conventional Cytology in a Randomized Setting. *J Natl Cancer Inst* 2009; 101(23):1612-1623.

Leinonen MK, Nieminen P, Lonnberg S, Malila N, Hakama M, Pokhrel A 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* 2012; 345:e7789.

Little J, Day N, Marteau T, Parmar M, Patnick J, Woodman C. Cytological surveillance compared with immediate referral for colposcopy in management of women with low grade cervical abnormalities: Multicentre randomised controlled trial. *BMJ (Online)* 2009; 339(7716):329.

Lonnberg S, Anttila A, Luostarinen T, Nieminen P. Age-specific effectiveness of the Finnish cervical cancer screening programme. *Cancer Epidemiol Biomarkers Prev* 2012; 21(8).

Lonnberg S, Nieminen P, Luostarinen T, Anttila A. Mortality audit of the Finnish cervical cancer screening program. *Int J Cancer* 2013; 132(9).

Lorenzato F, Singer A, Ho L, Santos L, Batista RdL, Lubambo T et al. Human papillomavirus detection for cervical cancer prevention with polymerase chain reaction in self-collected samples. Am J Obstet Gynecol 2002; 186(5):962-968.

Malila N, Leinonen M, Kotaniemi-Talonen L, Laurila P, Tarkkanen J, Hakama M. The HPV test has similar sensitivity but more overdiagnosis than the Pap testâ€"A randomised health services study on cervical cancer screening in Finland. *Int J Cancer* 2013; 132(9):2141-2147.

Mandelblatt J, Gopaul I, Wistreich M. Gynecological care of elderly women: Another look at papanicolaou smear testing. *JAMA* 1986; 256(3):367-371.

Marlow LA, Sangha A, Patnick J, Waller J. The Jade Goody Effect: whose cervical screening decisions were influenced by her story? *J Med Screen* 2012; 19(4):184-188.

McCaffery K, Irwig L, Chan S, Macaskill P, Barratt A, Lewicka M et al. HPV testing versus repeat Pap testing for the management of a minor abnormal Pap smear: Evaluation of a decision aid to support informed choice. *Patient education and counseling* 2008; 73(3):473-479.

McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol* 2008; 9(5).

Meijer CJLM, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer* 2009; 124(3):516-520.

Miller MG, Sung HY, Sawaya GF, Kearney KA, Kinney W, Hiatt RA. Screening interval and risk of invasive squamous cell cervical cancer. *Obstet Gynecol* 2003; 101(1).

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; 354(9193):1896-1900.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; 354(9193):1896-1900.

Moyer VA. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012; 156(12):880-91.

MSAC. Final Decision Analytic Protocol to guide the assessment of the National Cervical Screening Program Renewal. 2012. Canberra: Medical Services Advisory Committee.

MSAC. Human papillomavirus triage test for women with possible or definite low-grade squamous intra-epithelial lesions. Assessment Report 39. 2009. Canberra, Medical Services Advisory Committee.

Murphy J, Kennedy E, Dunn S, McLachlin C, Fung Kee Fung M, Gzik D et al. HPV testing in primary cervical screening: a systematic review and meta-analysis. J Obstet Gynaecol Can 2012; 34(5):443-52.

Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K et al. Human Papillomavirus and Papanicolaou Tests to Screen for Cervical Cancer. *N Engl J Med* 2007; 357(16):1589-1597.

Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K et al. Efficacy of HPV DNA Testing With Cytology Triage and/or Repeat HPV DNA Testing in Primary Cervical Cancer Screening. *J Natl Cancer Inst* 2009; 101(2):88-99.

NHMRC. A guide to the development, implementation and evaluation of clinical practice guidelines. 1999. Canberra, National Health and Medical Research Council.

NHMRC. How to use the evidence: assessment and application of scientific evidence. 2000. Canberra, National Health and Medical Research Council. Handbook series on preparing clinical practice guidelines.

NHMRC. NHMRC levels of evidence and grades for recommendations for developers of guidelines. 2009. Canberra, National Health and Medical Research Council, Commonwealth of Australia.

NHMRC. Screening to prevent cervical cancer: Guidelines for the management of asymptomatic women with screen detected abnormalities. 2005. Canberra, National Health and Medical Research Council.

Noehr B, Jensen A, Frederiksen K, Tabor A, Kjaer S. Loop electrosurgical excision of the cervix and subsequent risk for spontaneous preterm delivery: a population-based study of singleton deliveries during a 9-year period. *Am J Obstet Gynecol* 2009; 201(1):33e1-e6.

Noehr B, Tabor A, Frederiksen K, Kjaer S. Loop electrosurgical excision of the cervix and the subsequent risk of preterm delivery. *Acta Obstet Gynecol Scand* 2007; 86(5):596-603.

Ogilvie GS, Krajden M, van Niekerk DJ, Martin RE, Ehlen TG, Ceballos K et al. Primary cervical cancer screening with HPV testing compared with liquid-based cytology: results of round 1 of a randomised controlled trial - the HPV FOCAL Study. *Br J Cancer* 2012; 107(12):1917-1924.

Ogilvie GS, Niekerk DJ, Krajden M, Martin RE, Ehlen TG, Ceballos K et al. A randomized controlled trial of Human Papillomavirus (HPV) testing for cervical cancer screening: trial design and preliminary results (HPV FOCAL Trial). *BMC cancer* 2010; 10:111.

Ortoft G, Henriksen T, Hansen E, Petersen L. After conisation of the cervix, the perinatal mortality as a result of preterm delivery increases in subsequent pregnancy. *BJOG* 2010; 117(3):258-267.

Ostor AG. Natural history of cervical intraepithelial neoplasia: a critical review. *Int J Gynecol Pathol* 1993; 12(2).

Palmer TJ, Nicoll SM, Mckean ME, Park AJ, Bishop D, Baker L et al. Prospective parallel randomized trial of the MultiCyte(trademark) ThinPrep(registered trademark) imaging system: The Scottish experience. *Cytopathology* 2012.

Pantawala I, Bauer H, Miyamoto J, Park I, Huchko M, Smith-McCune K. A systematic review of randomized trials assessing human papillomavirus testing in cervical cancer screening. Am J Obstet Gynecol 2013; 208(5):343-53.

Paraskevaidis E, Koliopoulos G, Lolis E, Papanikou E, Malamou-Mitsi V, Agnantis N. Delivery outcomes following loop electrosurgical excision procedure for microinvasive (FIGO stage IA1) cervical cancer. *Gynecol Oncol* 2002; 86(1):10-13.

Patel A, Galaal K, Burnley C, Faulkner K, Martin-Hirsch P, Bland MJ et al. Cervical cancer incidence in young women: a historical and geographic controlled UK regional population study. *Br J Cancer* 2012; 106(11).

Peirson L, Fitzpatrick-Lewis D, Ciliska D, Warren R, Elit L. Screening for Cervical Cancer. 2012. Canadian Task Force on Preventive Health Care.
Ref Type: Report

Peto J, Gilham C, Deacon J, Taylor C, Evans C, Binns W et al. Cervical HPV infection and neoplasia in a large population-based prospective study: the Manchester cohort. *Br J Cancer* 2004a; 91(5).

Peto J, Gilham C, Fletcher O, Matthews FE. The cervical cancer epidemic that screening has prevented in the UK. *The Lancet* 2004b; 364(9430):249-256.

Petry KU, Menton S, Menton M, van Loenen-Frosch F, de Carvalho GH, Holz B et al. Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *Br J Cancer* 2003; 88(10):1570-1577.

Qiao Y, Sellors J, Eder P, Bao Y, Lim J, Zhao F et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. Lancet Oncol 2008; 9(10):929-936.

Queensland Government. Queensland Health Guidelines for the use of Liquid Based Cytology. Queensland Government [2012 [cited 2013 Aug. 23]; Available from: URL: http://www.health.qld.gov.au/qhpolicy/docs/gdl/qh-gdl-314.pdf

Rebolj M, van Ballegooijen M, Lynge E, Looman C, Essink-Bot ML, Boer R et al. Incidence of cervical cancer after several negative smear results by age 50: prospective observational study. *Brit Med J* 2009; 338.

Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkmans NWJ, Heideman DAM et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: Final results of the POBASCAM randomised controlled trial. *Lancet Oncol* 2012; 13(1):78-88.

Roberts JM, Thurloe JK, Bowditch RC, Hyne SG, Greenberg M, Clarke JM et al. A three-armed trial of the thinprep imaging system. *Diagn Cytopathol* 2007; 35(2):96-102.

Ronco G, Biggeri A, Confortini M, Naldoni C, Segnan N, Sideri M et al. Health technology assessment report: HPV DNA based primary screening for cervical cancer precursors. *Epidemiol Prev* 2012; 36(3-4 Suppl 1):e1-72.

Ronco G, Biggeri A, Confortini M, Naldoni C, Segnan N, Sideri M et al. [HPV DNA based primary screening for cervical cancer precursors]. Epidemiol Prev 2012; 36(3-4 Suppl 1):e1-e72.

Ronco G, Brezzi S, Carozzi F, Dalla Palma P, Giorgi-Rossi P, Minucci D et al. The New Technologies for Cervical Cancer Screening randomised controlled trial. An overview of results during the first phase of recruitment. *Gynecol Oncol* 2007; 107(1 SUPPL.):S230-S232.

Ronco G, Confortini M, Maccallini V, Naldoni C, Segnan N, Sideri M et al. [Use of liquid-based cytology for cervical cancer precursors screening]. Epidemiol Prev 2012; 36(5 Suppl 2):e1-e33.

Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Palma PD, Naldoni C et al. Accuracy of liquid based versus conventional cytology: Overall results of new technologies for cervical cancer screening: Randomised controlled trial. *Brit Med J* 2007a; 335(7609):28-31.

Ronco G, Cuzick J, Segnan N, Brezzi S, Carozzi F, Folicaldi S et al. HPV triage for low grade (L-SIL) cytology is appropriate for women over 35 in mass cervical cancer screening using liquid based cytology. *Eur J Cancer* 2007; 43(3):476-480.

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Palma PD, Del Mistro A et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* 2008; 100(7):492-501.

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Palma PD, Del Mistro A 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):249-257.

Ronco G, Segnan N, Giorgi-Rossi P, Zappa M, Casadei GP, Carozzi F et al. Human papillomavirus testing and liquid-based cytology: Results at recruitment from the new technologies for cervical cancer randomized controlled trial. *J Natl Cancer Inst* 2006; 98(11):765-774.

Ronco G, Segnan N. HPV testing for primary cervical cancer screening. *The Lancet* 2007; 370(9601):1740-1742.

Rossi PG, Carozzi F, Collina G, Confortini M, Dalla Palma P, De Lillo M et al. HPV testing is an efficient management choice for women with inadequate liquid-based cytology in cervical cancer screening. *Am J Clin Pathol* 2012; 138(1):65-71.

Sadler L, Saftlas A, Wang W, Exeter M, Whittaker J, McCowan L. Treatment for cervical intraepithelial neoplasia and risk of preterm delivery. *JAMA* 2004; 291(17):2100-2106.

Salmeron J, Lazcano-Ponce E, Lorincz A, Hernandez M, Hernandez P, Leyva A et al. Comparison of HPV-based assays with Papanicolaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes Control* 2003; 14(6).

Samson S, Bentley J, Fahey T, McKay D, Gill G. The effect of loop electrosurgical excision procedure on future pregnancy outcome. *Obstet Gynecol* 2005; 105(2):325-332.

Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS et al. A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. *Int J Cancer* 2005; 116(4):617-623.

Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM et al. HPV Screening for Cervical Cancer in Rural India. *N Engl J Med* 2009; 360(14):1385-1394.

Sanner K, Wikstrom I, Strand A, Lindell M, Wilander E. Self-sampling of the vaginal fluid at home combined with high-risk HPV testing. *Brit J Cancer* 2009; 101(5).

Sargent A, Bailey A, Almonte M, Turner A, Thomson C, Peto J et al. Prevalence of type-specific HPV infection by age and grade of cervical cytology: data from the ARTISTIC trial. *Br J Cancer* 2008; 98(10):1704-1709.

Sargent A, Bailey A, Turner A, Almonte M, Gilham C, Baysson H et al. Optimal Threshold for a Positive Hybrid Capture 2 Test for Detection of Human Papillomavirus: Data from the ARTISTIC Trial. *Journal of Clinical Microbiology* 2010; 48(2):554-558.

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).

Sasieni P, Castanon A, Cuzick J. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. *BMJ* 2009; 339.

Sasieni P, Castanon A. Dramatic increase in cervical cancer registrations in young women in 2009 in England unlikely to be due to the new policy not to screen women aged 20-24. *J Med Screen* 2012; 19(3):127-132.

Sasieni PD, Cuzick J, Lynch-Farmery E. Estimating the efficacy of screening by auditing smear histories of women with and without cervical cancer. The National Co-ordinating Network for Cervical Screening Working Group. *Br J Cancer* 1996; 73(8).

Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *Am J Clin Pathol* 2012; 137(4):516-542.

Sawaya GF, Kerlikowske K, Lee NC, Gildengorin G, Washington AE. Frequency of cervical smear abnormalities within 3 years of normal cytology. *Obstet Gynecol* 2000; 96(2):219-223.

Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study: Design, methods and characteristics of trial participants. *Acta Cytol* 2000; 44(5):726-742.

Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. The Lancet 370[9590], 890-907. 8-9-2007.

Schiffman M, Glass AG, Wentzensen N, Rush BB, Castle PE, Scott DR et al. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. *Cancer Epidemiol Biomarkers Prev* 2011; 20(7):1398-1409.

Scottish Cervical Cytology Review Group Feasibility Sub Group. Cervical Cytology Thinprep Imager® (TIS) Feasibility Study. Report from the Feasibility Sub Group to Cervical Cytology Review Group. 2009. Scottish Pathology Network.

Sherman ME, Schiffman M, Cox JT. Effects of age and human papilloma viral load on colposcopy triage: Data from the randomized atypical squamous cells of undetermined significance/low-grade squamous intraepithelial lesion triage study (ALTS). *J Natl Cancer Inst* 2002; 94(2):102-107.

Siebers AG, Klinkhamer PJJM, Grefte JMM, Massuger LFAG, Vedder JEM, Beijers-Broos A et al. Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: A randomized controlled trial. *J Am Med Assoc* 2009; 302(16):1757-1764.

Sigurdsson K, Sigvaldason H. Is it rational to start population-based cervical cancer screening at or soon after age 20? Analysis of time trends in preinvasive and invasive diseases. *Eur J Cancer* 2007; 43(4).

Sigurdsson K. Cervical cancer: cytological cervical screening in Iceland and implications of HPV vaccines. *Cytopathology* 2010; 21(4).

Snijders PJ, Verhoef VM, Arbyn M, Ogilvie G, Minozzi S, Banzi R et al. High-risk HPV testing on self-sampled versus clinician-collected specimens: a review on the clinical accuracy and impact on population attendance in cervical cancer screening. *Int J Cancer* 2013; 132(10).

Solomon D, Schiffman M, Tarone R. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: Baseline results from a randomized trial. *J Natl Cancer Inst* 2001; 93(4):293-299.

Stevens MW, Milne AJ, Parkinson IH, Nespolon WW, Fazzalari NL, Arora N et al. Effectiveness of AutoPap system location-guided screening in the evaluation of cervical cytology smears. *Diagn Cytopathol* 2004; 31(2):94-99.

Strander B, Andersson-Ellstrom A, Milsom I, Radberg T, Ryd W. Liquid-based cytology versus conventional Papanicolaou smear in an organized screening program : a prospective randomized study. *Cancer* 2007; 111(5):285-291.

Szarewski A, Cadman L, Mallett S, Austin J, Londesborough P, Waller J et al. Human papillomavirus testing by self-sampling: assessment of accuracy in an unsupervised clinical setting. *J Med Screen* 2007; 14(1).

Szarewski A, Cadman L, Mesher D, Austin J, Ashdown-Barr L, Edwards R et al. HPV self-sampling as an alternative strategy in non-attenders for cervical screening - a randomised controlled trial. *Br J Cancer* 2011; 104(6).

Tabrizi SN, Brotherton JML, Kaldor JM, Skinner SR, Cummins E, Liu B et al. Fall in Human Papillomavirus Prevalence Following a National Vaccination Program. *Journal of Infectious Diseases* 2012; 206(11):1645-1651.

Tamalet C, Le Retraite L, Leandri FX, Heid P, Sancho Garnier H, Piana L. Vaginal self-sampling is an adequate means of screening HR-HPV types in women not participating in regular cervical cancer screening. *Clin Microbiol Infect* 2013; 19(1).

Tan L, Pepra E, Haloob RK. The outcome of pregnancy after large loop excision of the transformation zone of the cervix. *J Obstet Gynaecol* 2004; 24(1):25-27.

Theraputic Goods Administration (TGA). Overview of the regulatory framework for in-vitro diagnostic medical devices. Australian Government, Department of Health and Ageing [2011 [cited 2013 Aug. 2]; Available from: URL: http://www.tga.gov.au/industry/ivd-framework-overview.htm

TOMBOLA (Trial of Management of Borderline and Other Low-grade Abnormal smears), Sharp L, Cotton S, Cochran C, Gray N, Little J et al. After-effects reported by women following colposcopy, cervical biopsies and LLETZ: results from the TOMBOLA trial. *BJOG* 2009; 116(11):1506-1514.

VCCR. Statistical Report. 2011. Victorian Cervical Cytology Registry.

Vesco KK, Whitlock EP, Eder M, Lin J, Burda BU, Senger CA et al. Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force. AHRQ Report No.: 11-05156-EF-1. Agency for Healthcare Research and Quality; 2011.

Virtanen A, Anttila A, Luostarinen T, Nieminen P. Self-sampling versus reminder letter: effects on cervical cancer screening attendance and coverage in Finland. *Int J Cancer* 2011a; 128(11).

Virtanen A, Nieminen P, Luostarinen T, Anttila A. Self-sample HPV Tests As an Intervention for Nonattendees of Cervical Cancer Screening in Finland: a Randomized Trial. *Cancer Epidem Biomar* 2011b; 20(9).

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003; 3(25).

Whitlock EP, Vesco KK, Eder M, Lin JS, Senger CA, Burda BU. Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2011; 155(10):687-215.

Wikstrom I, Lindell M, Sanner K, Wilander E. Self-sampling and HPV testing or ordinary Pap-smear in women not regularly attending screening: a randomised study. *Br J Cancer* 2011; 105(3).

Wilbur DC, Black-Schaffer WS, Luff RD, Abraham KP, Kemper C, Molina JT et al. The Becton Dickinson focalpoint GS imaging system: Clinical trials demonstrate significantly improved sensitivity for the detection of important cervical lesions. *Am J Clin Pathol* 2009; 132(5):767-775.

Woodman CB, Collins S, Winter H, Bailey A, Ellis J, Prior P et al. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. *Lancet* 2001; 357(9271).

Wright TJ, Denny L, Kuhn L, Pollack A, Lorincz A. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. JAMA 2000; 283(1):81-86.

Yang B, Morrell S, Zuo Y, Roder D, Tracey E, Jelfs P. A case-control study of the protective benefit of cervical screening against invasive cervical cancer in NSW women. *Cancer Causes Control* 2008; 19(6):569-576.

Zhao FH, Lin MJ, Chen F, Hu SY, Zhang R, Belinson JL et al. Performance of high-risk human papillomavirus DNA testing as a primary screen for cervical cancer: a pooled analysis of individual patient data from 17 population-based studies from China. *Lancet Oncol* 2010; 11(12):1160-1171.