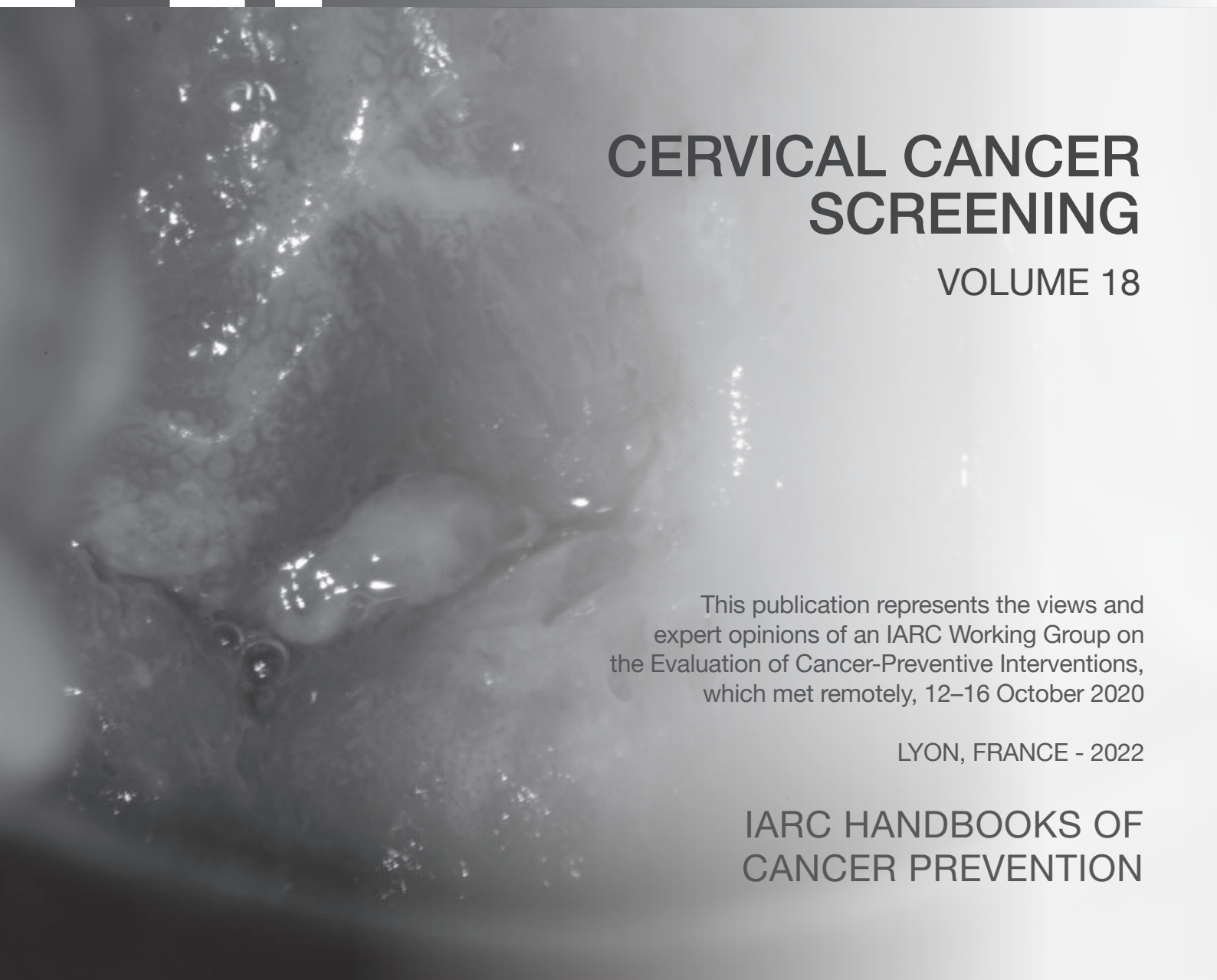
IARC HANDBOOKS



CERVICAL CANCER SCREENING

VOLUME 18

This publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Cancer-Preventive Interventions, which met remotely, 12–16 October 2020

LYON, FRANCE - 2022

IARC HANDBOOKS OF CANCER PREVENTION

International Agency for Research on Cancer

World Health Organization

4. PREVENTIVE AND ADVERSE EFFECTS OF CERVICAL CANCER SCREENING

4.1 Methodological issues

4.1.1 Considerations about beneficial effects of cervical screening

(a) General principles

This section considers the benefits of cervical screening, the accuracy of methods used for cervical screening and management, and the types of studies and data used to evaluate cervical screening and the related metrics to evaluate the benefits of screening.

The main goal of cervical screening is the prevention of invasive cervical cancer by the detection and treatment of intraepithelial precancer (see Section 1.2). This needs to be distinguished from downstaging, which is the early detection and treatment of already invasive cancer to improve the chance of a cure; downstaging is the main goal of screening for cancer types that lack well-defined, treatable precancerous precursors. Successful detection and treatment of precancers should lead to a reduction in cervical cancer incidence and mortality. Successful stage shift should lead to a reduction in cervical cancer mortality.

The theoretical maximum possible benefit of cervical screening in a population is the complete secondary prevention of invasive cancer by detecting and treating all cervical precancers that would progress to invasive cancer. The cumulative lifetime incidence of cervical cancer ranges from 1% to 5% of all women; for the other women, cervical cancer screening does not bring any benefits on a personal basis because they will never have the disease in any case, and thus it is essential to pay attention to its possible harms.

The use of cervical cancer screening with Pap cytology became widespread in many high-income countries during the late 1960s and the 1970s, before randomized trials became the standard for evaluating the efficacy of preventive interventions. Because of this, the initial evidence on the efficacy of cervical cancer screening was derived from ecological or surveillance data, cohort studies, and case–control studies (for details, see Section 4.3.2).

(b) Diagnostic accuracy

For a screening test to be accurate, it must, as a primary requirement, yield approximately the same result when repeated in the same and different test settings. Some tests are inherently subjective and often yield non-reproducible results in the case of minor cytological or minor visual abnormalities. Such tests are bound to be inaccurate.

Whatever type of cervical test is being evaluated, the same statistical analyses are applied to assess accuracy. Continuous or ordinal measurements (e.g. the viral load measured by a human papillomavirus [HPV] test or the

grades of cytological abnormality) are typically combined into a few categories before analysis (e.g. positive/negative or abnormal/normal). The accuracy of a screening test is measured as a trade-off between sensitivity and specificity, which are the well-known measures of test performance given outcome category (sensitivity is test positivity among precancers; specificity is test negativity given the absence of precancer or cancer). Sensitivity and specificity can be estimated with any major study design, including the common case–control study. An important derivative statistic that is based on sensitivity and specificity is the area under the curve (AUC) of a receiver operating characteristic (ROC) curve, which evaluates sensitivity and specificity over a wide range of cut-off values.

For the evaluation of screening tests, we distinguish between analytical accuracy and clinical accuracy. Analytical accuracy relates to the target of detection (e.g. HPV DNA), whereas clinical accuracy relates to the detection of cervical precancer. Achieving maximal analytical sensitivity is not the primary goal of cervical screening tests. HPV infection and its associated microscopic and visual abnormalities are common and are typically benign. The prevalence of HPV varies greatly by age and population and can be very high in some settings. A positive HPV test result (or low-grade squamous intraepithelial lesion [LSIL] cytology or visual impression of acetowhitening), which accurately detects infection with a carcinogenic HPV type, is, in the context of risk of precancer, a false-positive result, because most infections resolve or become undetectable without intervention. Unlike the situation for other infectious agents, considering all positive analytical test results to indicate a positive cervical screening result leads to poor specificity and low positive predictive value (PPV) in screening for cervical precancer. The challenge of cervical screening is to choose tests and thresholds that maximize accuracy for diagnosis of precancer as distinct from benign HPV effects.

Evaluating the accuracy of screening tests typically involves testing followed by the systematic application of the reference standard test, traditionally colposcopy-directed biopsy of all acetowhite lesions (Wentzensen et al., 2015), to all women enrolled in a relevant study population. All tests, including the reference standard, should be performed independently and within a very short time period. The principles and reporting standards for diagnostic accuracy studies are summarized by the Standards for Reporting of Diagnostic Accuracy Studies (STARD) criteria (Bossuyt et al., 2015). The quality of diagnostic accuracy studies included in a meta-analysis can be assessed by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist (Clarke et al., 2020).

It is not feasible or economically viable to apply the reference standard test to large populations of women attending cervical cancer screening; this would also be unethical, because it would result in a large number of women with a very low likelihood of having precancer undergoing colposcopy and biopsy. Clinical practice in cervical cancer screening usually involves a screening test, sometimes followed by triage; triage-positive or screen-positive women are referred for colposcopy and biopsy. Therefore, real-life screening data may suffer from partial and differential verification bias when absolute accuracy is estimated. When the screening test (e.g. visual inspection with acetic acid [VIA] or visual inspection with Lugol's iodine [VILI]) and the reference standard test (e.g. colposcopy) are subjective and correlated, this can lead to severely biased estimates (Arbyn et al., 2008a), unless intrinsic correlation is accounted for statistically (Leeflang & Reitsma, 2018). However, the risk of a cancer or even a precancer in women with a negative HPV test result is so low that it is not necessary to refer a fraction of HPV-negative women for further verification when a well-validated HPV

DNA test is used for primary screening. In fact, adjustment for verification bias in HPV-negative women can lead to substantial distortions in the estimates of test accuracy (Castle et al., 2020). Verification bias is usually a minor issue when relative accuracy (comparing one test directly with another) is assessed.

The design and evaluation of screening approaches depend on precise definition of the screening target. Precancer is the causal surrogate for cancer risk in this context; if defined formally, a reduction in precancer should translate into the same proportional reduction in cancer. However, there are no markers that accurately identify the lesions that would progress to cancer. Cervical cancer screening studies are usually based on cervical intraepithelial neoplasia grade 2 or worse (CIN2+) or CIN grade 3 or worse (CIN3+) as end-points. CIN3+ is a more reliable outcome, because this diagnosis is more reproducible and is more strongly associated with progression to cancer. If precancer is defined too broadly (e.g. including a subset of CIN2 caused by HPV types that are almost never found in cancer), tests evaluated against this inflated standard will have distorted evaluations. For example, an HPV test that correctly targets only the truly carcinogenic types would be incorrectly criticized for lack of sensitivity rather than being recognized for increased specificity. Although it is preferable to use CIN3/adenocarcinoma in situ (AIS) as the end-point in screening evaluations, treatment of CIN2 can lead to underestimation of risk of CIN3 because some of the treated CIN2 would progress to CIN3. This does not affect CIN3+ end-points in cross-sectional studies of previously unscreened individuals.

(c) Randomized screening trials

To study the health effects of a new screening technology in real-world screening settings, randomized screening trials have been conducted in several countries. Screening trials are pragmatic trials (Schwartz & Lellouch, 1967, 2009) embedded in routine screening with few additional inclusion criteria, and with realistic triage and management of surveillance. The intervention effect measured in such pragmatic trials will be close to the effect observed when implementing the new technology in the real world, and its interpretation will not be limited to the study trial.

The ultimate goal of cancer screening is the reduction of cancer mortality, but the effect on cancer mortality is very difficult to measure in countries with screening in place, and it has only been assessed in a previously unscreened population (Sankaranarayanan et al., 2009). The same limitations exist for the end-point cancer, which has only been studied in a pooled analysis of European screening trials (Ronco et al., 2014). Other screening trials have CIN3+ or CIN2+ as the primary end-point.

Randomized screening trials aim to directly estimate the effect of switching technology on the detection of CIN3+ and CIN2+ over one or two screening rounds. Results in the first round can also be studied by a prospective study, where the new and conventional technologies are used in parallel and women are managed on the basis of the results of all tests. However, the randomized trial and the combined testing design may give different results when the results from the new and conventional technologies are dependent for reasons unrelated to the development of cervical cancer. This is illustrated by two examples. The first example is a study in which primary HPV testing with cytological testing on HPV-positive samples is compared with cytology alone. The performance of cytology may be influenced by knowledge of the HPV status. A valid estimate of the effect of HPV testing on CIN3+ detection can be obtained through a randomized trial in which cytotechnicians in the intervention group are informed about the HPV status of the samples (Leinonen et al., 2012). The second example is a study in which liquid-based cytology is compared with conventional cytology. With the combined

testing design, the sampling procedure may place the second test at a disadvantage because cells for diagnosis have been removed by the first sample. This potential sampling effect can be avoided by randomizing women to one of the two test typesl (Ronco et al., 2007).

In most cervical cancer screening trials, participants are followed up for more than one screening round. Because the purpose of screening is to prevent cancer through the detection of precancer, the main aim of trials with CIN2+ or CIN3+ as end-points is to show that the new technology increases the lead-time gain from screening. This can be done by showing that increased detection of CIN2+ and CIN3+ in the first round of screening is followed by decreased detection in the second round.

Randomized screening trials vary in how they define the second round. Some trials categorize all CIN2+ and CIN3+ detected beyond a certain time point as in the second round (Rijkaart et al., 2012; Ogilvie et al., 2018). The strength of this approach is that a decreased detection of CIN2+ or CIN3+ in the second round can be explained by earlier detection, because randomization ensures that the risks of precancer at baseline are equal in the two study arms. The approach works well when the screening interval is long enough to ensure that all women in the comparison group have completed the first round. Otherwise, it may be better to select women for whom completion of the first round can be confirmed. To minimize the chance that the impact on lead-time gain is distorted by baseline differences between subgroups in the risk of precancer, the second round should include not only follow-up of women with a negative screening test result at baseline but also follow-up of screen-positive women with a negative test result at short-term repeat testing (Chan et al., 2020) and follow-up of women who underwent surveillance after colposcopy (Ronco et al., 2010).

Randomized screening trials also vary with respect to the choice of technology in the second round. Some trials use only the conventional technology in both arms (Naucler et al., 2007; Ronco et al., 2010), whereas others use the new technology in both arms (Rijkaart et al., 2012; Ogilvie et al., 2018) or retain separate screening strategies in the two arms (Kitchener et al., 2009). This may influence the trial results. For example, some of the precancers may remain undetected in women who are offered conventional technology in the first and second round, in particular when the difference in lead-time gain between the two technologies is large.

(d) Observational studies

Observational data play an important role in evaluating and improving cervical cancer screening programmes. Observational data range from ecological studies involving cancer registries to specific cohort studies that directly compare screening tests and strategies.

Cytology screening was introduced without evidence from randomized trials. Large decreases in the incidence of cervical cancer after the rapid implementation of cervical screening in some populations provided evidence of the effectiveness of cervical screening even from study designs that are typically not considered to be sufficient to prove causal associations between an intervention and a health effect. For example, large reductions in the incidence of cervical cancer were seen in Finland, Slovenia, and the United Kingdom after the implementation of national call–recall organized programmes (Quinn et al., 1999; Anttila, 2007; ZORA, 2018). Furthermore, the implementation of organized programmes in European countries, including Denmark, Finland, Italy, the Netherlands, Sweden, and the United Kingdom (Anttila et al., 2009), and integrated health systems in the USA, including Kaiser Permanente Northern California (Castle et al., 2018), as well as experiences with opportunistic screening in the Republic of Korea

(Odongua et al., 2007) and the USA (Landy et al., 2020), have led to the development of an infrastructure for the systematic collection of routine data on screening tests, results, and outcomes from screening and pathology registries. Many screening programmes continue to provide insights into the effectiveness of different screening protocols (Rebolj et al., 2008; Briët et al., 2010) and new technologies (Akamatsu et al., 2012; Rebolj et al., 2015; Rozemeijer et al., 2017; Forslund et al., 2019; Zorzi et al., 2020).

One of the strengths of these population-based studies is that, given the implementation of a programme that targets an entire population, it is possible to evaluate intention-to-treat approaches in cohort studies (e.g. Ronco et al., 2005), which reduces bias related to indication. When historical or geographical controls are used, the comparability of populations, even in the absence of indication bias, is a concern, particularly when two screened populations are compared to evaluate different protocols or technologies. In fact, two main determinants of cervical disease outcomes – the screening history (Maggino et al., 2016; Castle et al., 2019) and the prevalence of HPV (Bray et al., 2005; Sander et al., 2014) – can change rapidly over time and vary by geographical region.

Retrospective cohort studies examining sensitivity and efficacy against cancer have been used to compare screening tests; however, this study design has important methodological issues, which can lead to severely biased estimates when they are not properly accounted for. For example, studies that use a cancer diagnosis or the detection of a high-grade precursor lesion as a starting point and retrospectively select on previous screening results may be biased in favour of cytology when the management is differential between cytology and other tests and the screening history is limited (Blatt et al., 2015; Castle, 2015; Giorgi Rossi et al., 2016; Kaufman et al., 2020; Schiffman & Wentzensen, 2021). The choice of end-point is also meaningful in retrospective studies examining the performance of screening tests, and relates to the timing of previous testing. Most screening tests performed within a short time of cancer diagnosis are part of the clinical workup (Andrae et al., 2008; Castanon et al., 2013) or represent detection of an advanced, symptomatic cancer. This study design cannot capture the screening performance of these tests as an instrument to prevent cancer by detecting precancer (Ronco & Franceschi, 2018).

Well-designed observational studies have become important pillars of regulatory evaluations of cervical screening tests. For example, recent United States Food and Drug Administration (FDA) approvals of HPV tests for primary screening, either alone or in combination with cytology, were based not on randomized trials but on prospective cohort studies in which all comparator tests were conducted in the entire population and positive results from any test led to referral for colposcopy (FDA, 2019). These studies enable the efficient comparison of disease detection for different assays in the first screening round, but because the management is not differential for different test results, they do not enable the evaluation of disease outcomes by test result in subsequent screening rounds.

(e) Risk-based screening and management

Test sensitivity and specificity do not directly inform health decisions, which require knowledge of risk (i.e. the measures of outcome based on test result). Risk is measured over a defined time period (cross-sectional or, ideally, prospective). When population data are available, optimally cohort data from an observational study or trial, health decisions about screening can be made by answering practical questions about absolute risk: What is the (pre-test) risk of developing this cancer? (This informs whether screening is worth doing.) What is the risk of developing this cancer if the test result is positive, and what should be done next? How reassuring is a negative

test result, and when should a participant with a negative test result come back for another screen?

An accurate screening test will divide the population pre-test risk (i.e. the population prevalence of precancer) into substantially higher risk (PPV defined as a function of time from screening) when a test result is positive, or lower risk (1 – negative predictive value [NPV]) when a test result is negative. Risk stratification alone (i.e. the difference in post-test risk between those with a positive test result and those with a negative test result) is not meaningful without the context of clinical action thresholds. Meaningful risk stratification implies that the post-test risk for at least one of the groups (those with a positive test result or those with a negative test result) leads to different clinical management.

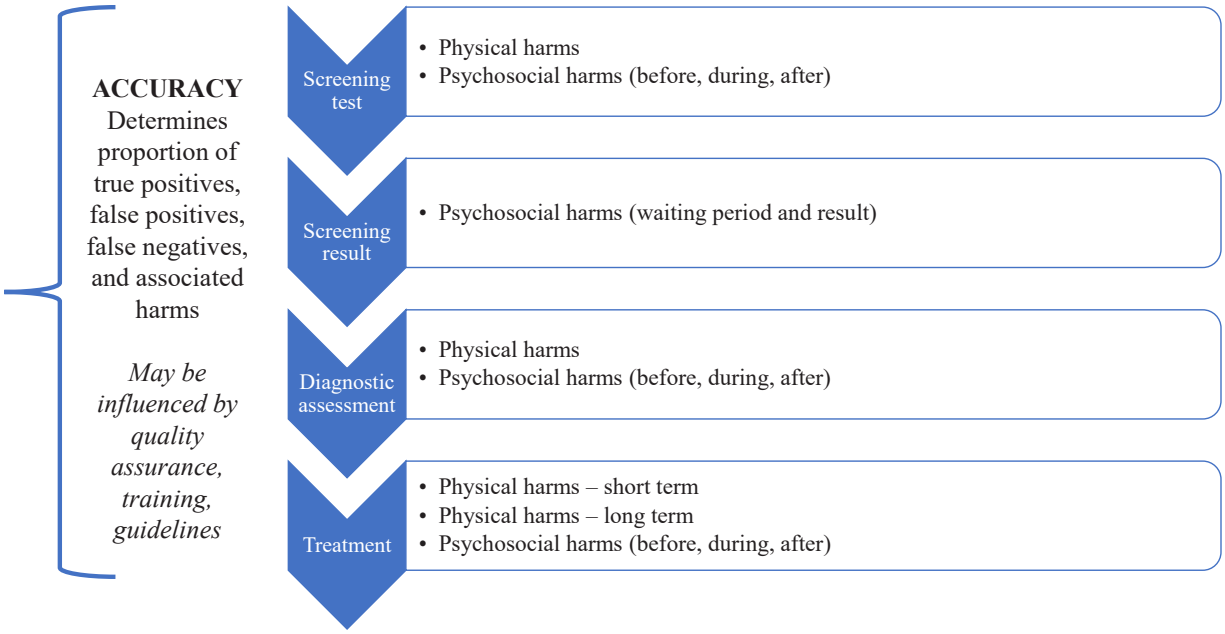
No single available cervical screening test has both very high PPV and very high NPV; therefore, a second, complementary triage test is generally used, which, in combination with the first test, provides a finer and more individually accurate level of risk discrimination. When the primary screening test is sensitive (e.g. in HPV testing), it is often reasonable to use the second test only to confirm the positive result from the first test, and to save the resources that would be required to co-test everyone. The combined results of screening and triage tests are grouped into categories, and the sensitivity/specificity or predictive values/risks of the combined strategy are assessed similarly as for a single test.

The same approach applies to screening, triage, post-colposcopy management, and post-treatment management. A risk-based approach may enable practice to be unified independent of the underlying tests. The 2019 update of the consensus guidelines for management of cervical cancer screening abnormalities (Perkins et al., 2020) adopted this principle as the foundation of the clinical guidelines. It is important to evaluate whether absolute risk estimates are portable between different populations. Even if the risk estimates apply across different populations, the decision thresholds may be adapted to clinical and societal preferences in different settings.

4.1.2 Considerations about harms of cervical screening

All cancer screening programmes involve potential harms, which individuals must balance against the potential benefits in deciding whether to participate in screening. Potential physical and psychological harms are considered in detail for each screening intervention or diagnostic step reviewed in this *Handbook*. Social and economic harms are generally not considered. Physical harms (e.g. pain, bleeding, and discharge) include those experienced because of the application of the initial screening test, as a consequence of follow-up, confirmatory, or diagnostic tests for women who receive a positive test result, or during or after treatment for screen-detected lesions. Psychological harms (e.g. anxiety and distress) may occur before, during, or after screening and may relate to the screening experience itself or to the receipt of the results and the perceived implications for the individual who has undergone a screening test, diagnostic test, or treatment procedures. Some harms, for example those that occur because of a false-positive test result, come about as a result of test characteristics or the screening system itself, and may not be observable directly by women or their clinicians. These harms may have effects at the population level; for example, false-positive screening test results may lead to unnecessary examinations and treatments, which, consequently, cause harm to women and waste medical resources. When policy-makers decide whether to implement a population-based screening programme, they must explicitly weigh the balance of potential benefits against potential harms at the population level (see Section 2.3). Fig. 4.1 presents a schematic overview of the potential harms associated with the cervical screening pathway.

Fig. 4.1 Potential harms associated with the cervical screening pathway



Created by the Working Group.

Harms pertaining to any screening technique are presented in this section. Evidence relating to potential harms specific to a technique, including their nature and rates of occurrence as observed during screening, is provided by technique in the relevant sections of this *Handbook* for screening by visual inspection (see Section 4.2.3), cytology (see Section 4.3.5), HPV testing (see Section 4.4.8), colposcopy (see Section 4.5), and treatment (see Section 1.2.5).

Ideally, a screening test to be used in a population will have a high NPV, which enables most women at risk of cervical cancer to be identified and the women with a negative test result to be correctly reassured that they are at low risk until the next screening test is due. The number of women potentially harmed can be measured as 1/PPV, which is the number of positive screening test results needed to confirm one precancer. Because of the natural history of HPV infection and disease (see Sections 1.2.1 and 1.2.2), the choice of screening interval, as well as the specificity of the test itself, will influence the rate of false-positive test results. Given the transient nature of most HPV infections, screening very frequently, either for HPV infection or for the cellular or visual changes associated with it, will be more likely to identify acute infection or disease with no potential for malignancy, thus increasing the proportion and number of false-positive test results and the potential harms.

Some of the concepts relevant to the monitoring of harms in cervical cancer screening programmes are discussed here.

(a) Overscreening

Cervical cancer screening that is carried out more frequently than is recommended in the current guidelines or that is used in a wider target age range or after hysterectomy can be called overscreening. The results from a decision analysis suggested that a short screening

Table 4.1 Harms associated with cervical cancer screening and management of screen-positive women in the USA and the Netherlands in 2007[a]

Event	Events per 1000 women		USA:Netherlands ratio
	USA	Netherlands	
Pap test			
Number	394	164	2.4
Symptoms[b] for at least 2–7 days	51	21	2.4
Abnormal test results			
Number	25	9	2.8
Anxiety for at least 12 weeks	9	3	3.0
Punch biopsy			
Number	16.9	4.3	3.9
Light[c] symptoms	19	5	3.8
Moderate or strong[c] symptoms	11	3	3.7
Treatment			
Number	3.0	1.8	1.7
Light[c] symptoms	2.4	1.4	1.7
Moderate or strong[c] symptoms	3.8	2.3	1.7

[a] Data were standardized to the female population of the USA aged 21–65 years in 2007.
[b] Lower abdominal pain, urinary discomfort, feeling sick, feeling dizzy, and/or painful sexual activity.
[c] Light refers to very light or light pain, bleeding, or discharge; moderate or strong refers to moderate, severe, or very severe pain, bleeding, or discharge.
Reproduced from Habbema et al. (2017). Copyright 2017, John Wiley & Sons.

interval and the use of HPV screening in women younger than 30 years will lead to an increase in the number of unnecessary colposcopies (Kim et al., 2018). Based on a systematic review, the main types of overscreening are screening that is too frequent (more frequent than the guideline recommendation), screening after hysterectomy, screening started before the recommended age, and screening after the recommended age at which screening should be stopped (Alber et al., 2018; Kim et al., 2018). A study in France reviewed outcomes for 63 821 women aged 25–65 years screened for up to 9 years, 37% of whom underwent cervical cancer screening at the recommended interval (every 3 years) and 63% more frequently. Overscreened women were more than twice as likely to have a CIN1 lesion diagnosed (age-adjusted relative risk, 2.09; 95% confidence interval [CI], 1.76–2.51) (Thiery et al., 2017).

Before the introduction of HPV testing, different screening intervals and ages were recommended in the USA and the Netherlands. Habbema et al. (2017) studied harms associated with cervical cancer screening and management of screen-positive women in the USA and the Netherlands. They included data on the number of Pap tests, abnormal test results, punch biopsies, treatments, and adverse effects of treatment (Table 4.1). The more intensive screening in the USA led to substantially higher rates of harms, with similar effects of screening on cervical cancer incidence and mortality in the two countries.

(b) Overdiagnosis and overtreatment

The target lesion for detection in cervical screening is the precursor lesions (high-grade squamous intraepithelial lesion [HSIL]/AIS), and the preventive effect on cervical cancer incidence is through treatment of these lesions.

Downstaging of cervical cancers discovered via screening is a secondary benefit that may also contribute to reductions in cervical cancer mortality achieved through screening (see Section 4.1.1). Overdiagnosis is defined as the diagnosis of a cancer as a result of screening that would not have been diagnosed in the patient's lifetime if screening had not taken place. Harms related to overdiagnosis are caused both by the physical harms associated with treatment and by the psychosocial consequences of a cancer diagnosis. Some authors have argued that because not all CIN3 lesions will result in cancer in a woman's lifetime if left untreated, the diagnosis of CIN3 itself should be described as overdiagnosis (Malila et al., 2013; Hakama et al., 2015; van Luijt et al., 2016). However, because a significant proportion of CIN3 lesions will progress to invasive cancer (Braun et al., 2011) and it is not possible to know which lesions may be safely left untreated, the use of the term overdiagnosis in this context might have unintended effects and lead to a reduction in the treatment of women with CIN3 lesions, followed by a concomitant rise in cervical cancer rates (Paul et al., 2018).

Overtreatment is defined as the treatment of a lesion that would never have progressed to be clinically recognized during a woman's lifetime. In relation to cervical screening, precancerous lesions are asymptomatic and are only detected through screening or incidentally in the investigation of other gynaecological conditions. In cervical cancer screening, there is a potential for overtreatment because of false-positive results, misdiagnosis, and conservative overclassification of histopathology of a lower grade. Overtreatment also occurs when lesions with no malignant potential are identified as precancers (HSIL/AIS) that require treatment. HSIL encompasses both CIN2 and CIN3. Whereas CIN3 reliably represents transforming infection with malignant potential, CIN2 includes a mixture of lesions that indicate both florid productive infection and true transforming infection. Because of an inability to reliably distinguish CIN2 lesions with true malignant potential from other lesions, most guidelines have recommended that CIN2 is also included as a treatment target (Arbyn et al., 2008b; Saslow et al., 2012; WHO, 2013; Jeronimo et al., 2016). However, the likelihood of progression from CIN2 to invasive cancer is lower than that of CIN3. The clinical course of untreated CIN2 at 24 months is 50% regression, 32% persistence, and 18% progression to CIN3 (Tainio et al., 2018). Methods of refining the diagnosis of CIN2 lesions so that the potential for progression can be better understood (e.g. through genotyping or molecular markers) may be strategies to reduce overtreatment. Age may also be a significant predictor of the likelihood of HSIL regression, because older women are less likely to experience regression of screen-detected lesions (Bekos et al., 2018). As described in Sections 1.2.1 and 1.2.2, the likelihood of a given lesion progressing to cancer will also be influenced by factors such as the causal HPV type, the woman's HIV status, and immunosuppression.

Overtreatment of CIN at grades below the accepted treatment thresholds may occur after referral due to abnormal cytology as part of the diagnostic process (e.g. via cone biopsy) or despite the availability of treatment guidelines (Volante et al., 2012; Nowakowski et al., 2016; Aitken et al., 2019).

References

Aitken CA, Siebers AG, Matthijsse SM, Jansen EEL, Bekkers RLM, Becker JH, et al. (2019). Management and treatment of cervical intraepithelial neoplasia in the Netherlands after referral for colposcopy. *Acta Obstet Gynecol Scand*. 98(6):737–46. doi:10.1111/aogs.13547 PMID:30687935

Akamatsu S, Kodama S, Himeji Y, Ikuta N, Shimagaki N (2012). A comparison of liquid-based cytology with conventional cytology in cervical cancer screening. *Acta Cytol*. 56(4):370–4. doi:10.1159/000337641 PMID:22846595

Alber JM, Brewer NT, Melvin C, Yackle A, Smith JS, Ko LK, et al. (2018). Reducing overuse of cervical cancer screening: a systematic review. *Prev Med*. 116:51–9. doi:10.1016/j.ypmed.2018.08.027 PMID:30149037

Andrae B, Kemetli L, Sparén P, Silfverdal L, Strander B, Ryd W, et al. (2008). Screening-preventable cervical cancer risks: evidence from a nationwide audit in Sweden. *J Natl Cancer Inst*. 100(9):622–9. doi:10.1093/jnci/djn099 PMID:18445828

Anttila A (2007). Cervical cancer screening is effective – the Finnish experience. *Entre Nous*. 64:26–8. Available from: https://www.euro.who.int/__data/assets/pdf_file/0007/69766/en64.pdf.

Anttila A, von Karsa L, Aasmaa A, Fender M, Patnick J, Rebolj M, et al. (2009). Cervical cancer screening policies and coverage in Europe. *Eur J Cancer*. 45(15):2649–58. doi:10.1016/j.ejca.2009.07.020 PMID:19699081

Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al., editors (2008b). European guidelines for quality assurance in cervical cancer screening. 2nd ed. Luxembourg: Office for Official Publications of the European Communities.

Arbyn M, Sankaranarayanan R, Muwonge R, Keita N, Dolo A, Mbalawa CG, et al. (2008a). Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *Int J Cancer*. 123(1):153–60. doi:10.1002/ijc.23489 PMID:18404671

Bekos C, Schwameis R, Heinze G, Gärner M, Grimm C, Joura E, et al. (2018). Influence of age on histologic outcome of cervical intraepithelial neoplasia during observational management: results from large cohort, systematic review, meta-analysis. *Sci Rep*. 8(1):6383. doi:10.1038/s41598-018-24882-2 PMID:29686397

Blatt AJ, Kennedy R, Luff RD, Austin RM, Rabin DS (2015). Comparison of cervical cancer screening results among 256,648 women in multiple clinical practices. *Cancer Cytopathol*. 123(5):282–8. doi:10.1002/cncy.21544 PMID:25864682

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al.; STARD Group (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 351:h5527. doi:10.1136/bmj.h5527 PMID:26511519

Braun S, Reimer D, Strobl I, Wieland U, Wiesbauer P, Müller-Holzner E, et al. (2011). Fatal invasive cervical cancer secondary to untreated cervical dysplasia: a case report. *J Med Case Reports*. 5(1):316. doi:10.1186/1752-1947-5-316 PMID:21767367

Bray F, Loos AH, McCarron P, Weiderpass E, Arbyn M, Møller H, et al. (2005). Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer Epidemiol Biomarkers Prev*. 14(3):677–86. doi:10.1158/1055-9965.EPI-04-0569 PMID:15767349

Briët MC, Berger TH, van Ballegooijen M, Boon ME, Rebolj M (2010). Effects of streamlining cervical cancer screening the Dutch way: consequences of changes in the Dutch KOPAC-based follow-up protocol and consensus-based limitation of equivocal cytology. *Acta Cytol*. 54(6):1095–100. doi:10.1159/000325251 PMID:21428155

Castanon A, Leung VMW, Landy R, Lim AWW, Sasieni P (2013). Characteristics and screening history of women diagnosed with cervical cancer aged 20–29 years. *Br J Cancer*. 109(1):35–41. doi:10.1038/bjc.2013.322 PMID:23820257

Castle PE (2015). Comparison of cervical cancer screening results among 256,648 women in multiple clinical practices. *Cancer Cytopathol*. 123(9):566. doi:10.1002/cncy.21572 PMID:26179787

Castle PE, Kinney WK, Xue X, Cheung LC, Gage JC, Poitras NE, et al. (2019). Role of screening history in clinical meaning and optimal management of positive cervical screening results. *J Natl Cancer Inst*. 111(8):820–7. doi:10.1093/jnci/djy192 PMID:30576462

Castle PE, Kinney WK, Xue X, Cheung LC, Gage JC, Zhao FH, et al. (2018). Effect of several negative rounds of human papillomavirus and cytology co-testing on safety against cervical cancer: an observational cohort study. *Ann Intern Med*. 168(1):20–9. doi:10.7326/M17-1609 PMID:29181509

Castle PE, Pierz AJ, Adcock R, Aslam S, Basu PS, Belinson JL, et al. (2020). A pooled analysis to compare the clinical characteristics of human papillomavirus-positive and -negative cervical precancers. *Cancer Prev Res (Phila)*. 13(10):829–40. doi:10.1158/1940-6207.CAPR-20-0182 PMID:32655005

Chan KKL, Liu SS, Wei N, Ngu SF, Chu MMY, Tse KY, et al. (2020). Primary HPV testing with cytology versus cytology alone in cervical screening – a prospective randomized controlled trial with two rounds of screening in a Chinese population. *Int J Cancer*. 147(4):1152–62. doi:10.1002/ijc.32861 PMID:31922265

Clarke MA, Darragh TM, Nelson E, Unger ER, Zuna R, Cremer M, et al. (2020). Reporting and assessing the quality of diagnostic accuracy studies for cervical cancer screening and management. *J Low Genit Tract Dis*. 24(2):157–66. doi:10.1097/LGT.0000000000000527 PMID:32243311

FDA (2019). FDA executive summary. New approaches in the evaluation for high-risk human papillomavirus nucleic acid detection devices. Prepared for the March 8, 2019 meeting of the Microbiology Devices Panel of the Medical Devices Advisory Committee. Silver Spring (MD), USA: United States Food and Drug Administration. Available from: https://www.fda.gov/media/122799/download.

Forslund O, Elfström KM, Lamin H, Dillner J (2019). HPV-mRNA and HPV-DNA detection in samples taken up to seven years before severe dysplasia of cervix

uteri. *Int J Cancer*. 144(5):1073–81. doi:10.1002/ijc.31819 PMID:30125346

Giorgi Rossi P, Ronco G, Dillner J, Elfström KM, Snijders PJF, Arbyn M, et al.; New Technologies for Cervical Cancer Group (2016). Why follow-back studies should be interpreted cautiously: the case of an HPV-negative cervical lesion. *Cancer Cytopathol*. 124(1):66–7. doi:10.1002/cncy.21622 PMID:26356132

Habbema D, Weinmann S, Arbyn M, Kamineni A, Williams AE, de Kok IMCM, et al. (2017). Harms of cervical cancer screening in the United States and the Netherlands. *Int J Cancer*. 140(5):1215–22. doi:10.1002/ijc.30524 PMID:27864938

Hakama M, Pokhrel A, Malila N, Hakulinen T (2015). Sensitivity, effect and overdiagnosis in screening for cancers with detectable pre-invasive phase. *Int J Cancer*. 136(4):928–35. doi:10.1002/ijc.29053 PMID:24975995

Jeronimo J, Castle PE, Temin S, Denny L, Gupta V, Kim JJ, et al. (2016). Secondary prevention of cervical cancer: ASCO resource-stratified clinical practice guideline. *J Glob Oncol*. 3(5):635–57. doi:10.1200/JGO.2016.006577 PMID:29094101

Kaufman HW, Alagia DP, Chen Z, Onisko A, Austin RM (2020). Contributions of liquid-based (Papanicolaou) cytology and human papillomavirus testing in cotesting for detection of cervical cancer and precancer in the United States. *Am J Clin Pathol*. 154(4):510–6. doi:10.1093/ajcp/aqaa074 PMID:32637991

Kim JJ, Burger EA, Regan C, Sy S (2018). Screening for cervical cancer in primary care: a decision analysis for the US Preventive Services Task Force. *JAMA*. 320(7):706–14. doi:10.1001/jama.2017.19872 PMID:30140882

Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, et al. (2009). HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol*. 10(7):672–82. doi:10.1016/S1470-2045(09)70156-1 PMID:19540162

Landy R, Mathews C, Robertson M, Wiggins CL, McDonald YJ, Goldberg DW, et al. (2020). A statewide population-based evaluation of cervical cancers arising during opportunistic screening in the United States. *Gynecol Oncol*. 159(2):344–53. doi:10.1016/j.ygyno.2020.08.033 PMID:32977987

Leeflang MMG, Reitsma JB (2018). Systematic reviews and meta-analyses addressing comparative test accuracy questions. *Diagn Progn Res*. 2(1):17. doi:10.1186/s41512-018-0039-0 PMID:31093565

Leinonen MK, Nieminen P, Lönnberg S, Malila N, Hakama M, Pokhrel A, et al. (2012). Detection rates of precancerous and cancerous cervical lesions within one screening round of primary human papillomavirus DNA testing: prospective randomised trial in Finland. *BMJ*. 345:e7789. doi:10.1136/bmj.e7789 PMID:23197596

Maggino T, Sciarrone R, Murer B, Dei Rossi MR, Fedato C, Maran M, et al. (2016). Screening women for cervical cancer carcinoma with a HPV mRNA test: first results from the Venice pilot program. *Br J Cancer*. 115(5):525–32. doi:10.1038/bjc.2016.216 PMID:27490801

Malila N, Leinonen M, Kotaniemi-Talonen L, Laurila P, Tarkkanen J, Hakama M (2013). 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*. 132(9):2141–7. doi:10.1002/ijc.27850 PMID:22987601

Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K, et al. (2007). Human papillomavirus and Papanicolaou tests to screen for cervical cancer. [Erratum in: N Engl J Med. 2008 359(15):1637. Johansson, Bo (added)] *N Engl J Med*. 357(16):1589–97. doi:10.1056/NEJMoa073204 PMID:17942872

Nowakowski A, Śliwczyński A, Seroczyński P, Cybulski M, Teter Z (2016). Reimbursed costs of management of uterine cervical lesions in Poland – a descriptive analysis of data from the National Health Fund and the Ministry of Health. *Cent Eur J Public Health*. 24(2):163–8. doi:10.21101/cejph.a4737 PMID:27178030

Odongua N, Chae YM, Kim MR, Yun JE, Jee SH (2007). Associations between smoking, screening, and death caused by cervical cancer in Korean women. *Yonsei Med J*. 48(2):192–200. doi:10.3349/ymj.2007.48.2.192 PMID:17461516

Ogilvie GS, van Niekerk D, Krajden M, Smith LW, Cook D, Gondara L, et al. (2018). Effect of screening with primary cervical HPV testing vs cytology testing on high-grade cervical intraepithelial neoplasia at 48 months: the HPV FOCAL randomized clinical trial. *JAMA*. 320(1):43–52. doi:10.1001/jama.2018.7464 PMID:29971397

Paul C, Sharples KJ, Baranyai J, Jones RW, Skegg DC (2018). Outcomes for women without conventional treatment for stage 1A (microinvasive) carcinoma of the cervix. *Aust N Z J Obstet Gynaecol*. 58(3):321–9. doi:10.1111/ajo.12753 PMID:29405269

Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al.; 2019 ASCCP Risk-Based Management Consensus Guidelines Committee (2020). 2019 ASCCP Risk-Based Management Consensus Guidelines for abnormal cervical cancer screening tests and cancer precursors. [Erratum in: J Low Genit Tract Dis. 2020 24(4):427] *J Low Genit Tract Dis*. 24(2):102–31. doi:10.1097/LGT.0000000000000525 PMID:32243307

Quinn M, Babb P, Jones J, Allen E; United Kingdom Association of Cancer Registries (1999). Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. *BMJ*. 318(7188):904–8. doi:10.1136/bmj.318.7188.904 PMID:10102852

Rebolj M, Rask J, van Ballegooijen M, Kirschner B, Rozemeijer K, Bonde J, et al. (2015). Cervical histology after routine ThinPrep or SurePath liquid-based cytology and computer-assisted reading in Denmark. *Br J Cancer*. 113(9):1259–74. doi:10.1038/bjc.2015.339 PMID:26448176

Rebolj M, van Ballegooijen M, van Kemenade F, Looman C, Boer R, Habbema JDF (2008). No increased risk for cervical cancer after a broader definition of a negative Pap smear. *Int J Cancer*. 123(11):2632–5. doi:10.1002/ijc.23803 PMID:18767046

Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkmans NWJ, Heideman DAM, et al. (2012). Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol*. 13(1):78–88. doi:10.1016/S1470-2045(11)70296-0 PMID:22177579

Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Dalla Palma P, Naldoni C, et al. (2007). Accuracy of liquid based versus conventional cytology: overall results of New Technologies for Cervical Cancer Screening randomised controlled trial. *BMJ*. 335(7609):28. doi:10.1136/bmj.39196.740995.BE PMID:17517761

Ronco G, Dillner J, Elfström KM, Tunesi S, Snijders PJF, Arbyn M, et al.; International HPV screening working group (2014). Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet*. 383(9916):524–32. doi:10.1016/S0140-6736(13)62218-7 PMID:24192252

Ronco G, Franceschi S (2018). Cervical cancer screening: the transformational role of routine human papillomavirus testing. *Ann Intern Med*. 168(1):75–6. doi:10.7326/M17-2872 PMID:29181508

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al.; New Technologies for Cervical Cancer screening (NTCC) Working Group (2010). Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol*. 11(3):249–57. doi:10.1016/S1470-2045(09)70360-2 PMID:20089449

Ronco G, Pilutti S, Patriarca S, Montanari G, Ghiringhello B, Volante R, et al.; Turin Cervical Screening Working Group (2005). Impact of the introduction of organised screening for cervical cancer in Turin, Italy: cancer incidence by screening history 1992–98. *Br J Cancer*. 93(3):376–8. doi:10.1038/sj.bjc.6602705 PMID:16012518

Rozemeijer K, Naber SK, Penning C, Overbeek LIH, Looman CWN, de Kok IMCM, et al. (2017). Cervical cancer incidence after normal cytological sample in routine screening using SurePath, ThinPrep, and conventional cytology: population based study. *BMJ*. 356:j504. doi:10.1136/bmj.j504 PMID:28196844

Sander BB, Rebolj M, Lynge E (2014). Trends of cervical cancer in Greenland: a 60-year overview. *Acta Oncol*. 53(4):452–61. doi:10.3109/0284186X.2014.883462 PMID:24512341

Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. (2009). HPV screening for cervical cancer in rural India. *N Engl J Med*. 360(14):1385–94. doi:10.1056/NEJMoa0808516 PMID:19339719

Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al.; ACS-ASCCP-ASCP Cervical Cancer Guideline 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. *CA Cancer J Clin*. 62(3):147–72. doi:10.3322/caac.21139 PMID:22422631

Schiffman M, Wentzensen N (2021). Cervical screening performance. *Am J Clin Pathol*. 155(4):616–20. doi:10.1093/ajcp/aqaa198 PMID:33399187

Schwartz D, Lellouch J (1967). Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis*. 20(8):637–48. doi:10.1016/0021-9681(67)90041-0 PMID:4860352

Schwartz D, Lellouch J (2009). Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol*. 62(5):499–505. doi:10.1016/j.jclinepi.2009.01.012 PMID:19348976

Tainio K, Athanasiou A, Tikkinen KAO, Aaltonen R, Cárdenas Hernándes J, Glazer-Livson S, et al. (2018). Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis. *BMJ*. 360:k499. doi:10.1136/bmj.k499 PMID:29487049

Thiery A, Akladios C, Fender M, Severac F, Baldauf JJ (2017). Excess cervical cancer screening smears: any benefit? A retrospective cohort in Alsace, France. *J Med Screen*. 24(2):92–7. doi:10.1177/0969141316647091 PMID:27206814

van Luijt PA, Rozemeijer K, Naber SK, Heijnsdijk EAM, van Rosmalen J, van Ballegooijen M, et al. (2016). The role of pre-invasive disease in overdiagnosis: a microsimulation study comparing mass screening for breast cancer and cervical cancer. *J Med Screen*. 23(4):210–6. doi:10.1177/0969141316629505 PMID:27068430

Volante R, Giubilato P, Ronco G (2012). Cervical cancer screening in Italy: quality of colposcopy and treatment. 2010 activity. [in Italian] *Epidemiol Prev*. 36(6 Suppl 1):78–86. PMID:23293272

Wentzensen N, Walker JL, Gold MA, Smith KM, Zuna RE, Mathews C, et al. (2015). Multiple biopsies and detection of cervical cancer precursors at colposcopy. *J Clin Oncol*. 33(1):83–9. doi:10.1200/JCO.2014.55.9948 PMID:25422481

WHO (2013). WHO guidelines for screening and treatment of precancerous lesions for cervical cancer prevention. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/94830; https://www.who.int/reproductivehealth/publications/cancers/screening_and_treatment_of_precancerous_lesions/en/.

ZORA (2018). National Cervical Cancer Screening Programme ZORA. Ljubljana, Slovenia: Cervical Cancer Screening Program and Registry ZORA. [in Slovenian] Available from: https://zora.onko-i.si/publikacije/kazalniki/.

Zorzi M, Del Mistro A, Giorgi Rossi P, Laurino L, Battagello J, Lorio M, et al. (2020). Risk of CIN2 or more severe lesions after negative HPV-mRNA E6/E7 overexpression assay and after negative HPV-DNA test: concurrent cohorts with a 5-year follow-up. *Int J Cancer*. 146(11):3114–23. doi:10.1002/ijc.32695 PMID:31583683

4.2 Screening by visual inspection

4.2.1 Visual inspection techniques

Visual techniques used in cervical screening include naked-eye examination with acetic acid (VIA) or Lugol's iodine (VILI) and camera-enhanced visual inspection. Naked-eye examination (with VIA or VILI) is a simple test for the early detection of cervical precancerous lesions and early invasive cancer and has been widely used in low- and middle-income countries (LMICs) to screen women for cervical precancer (Sankaranarayanan et al., 1998; Sankaranarayanan & Wesley, 2003). More recently, camera-enhanced image capture has been used to improve the performance of VIA (e.g. digital cervicography, smartphone attachments, intravaginal endoscopes, and portable monoscopic devices) (Parham et al., 2015; Goldstein et al., 2020; Xue et al., 2020) (see also Section 4.6.1). To date, no large randomized controlled trials (RCTs) have been performed that would enable the objective assessment of the effectiveness of enhanced VIA systems to detect precancer compared with routine VIA.

Recently, the combination of related novel technologies has enabled the development of artificial intelligence (AI) devices, which may supersede current technologies (see Section 4.6.1).

(a) Description of procedures

Visual inspection is appropriate for use in women in whom the squamocolumnar junction (SCJ) is visible (typically those younger than 50 years). In older, postmenopausal women, the SCJ gradually recedes into the endocervical canal, and it is possible to miss lesions when relying on visual inspection. Similarly, visual inspection cannot be used in younger women with a type 3 transformation zone (TZ). Therefore, before visual inspection is performed, the TZ type first needs to be accurately assessed (see Section 1.2.5, Fig. 1.18).

(i) Visual inspection with acetic acid (VIA)

Acetic acid causes dehydration of the cells of the cervical epithelium and some surface coagulation of cellular proteins, which reduces the transparency of the epithelium. These changes are more pronounced in abnormal epithelium, because of the higher nuclear density and consequent high concentration of proteins (Sankaranarayanan et al., 1998). After the application of acetic acid, more light is reflected back, making the epithelium appear white. The cervix is viewed with the naked eye through a vaginal speculum with the patient in either the left lateral position (dorsal with legs flexed) or the lithotomy position. VIA requires a good light source and freshly prepared 3–5% acetic acid in distilled water, and the examination should be carried out by a trained health-care provider (Sankaranarayanan & Wesley, 2003; WHO, 2014).

After gently removing any mucus from the cervix, the provider applies the acetic acid solution using a soaked swab or a spray bottle, and then looks to see if any white changes appear. The results of VIA examination are categorized as negative, positive, or suspicious for cancer (Table 4.2; Sankaranarayanan & Wesley, 2003; WHO, 2017). Acetowhite changes on the cervix that do not recede after 1 minute are likely to be associated with cervical precancer or cancer. If these changes are seen in the TZ and have well-defined borders, it is considered a positive result (WHO, 2013a, 2014). A positive VIA test result will reveal an area or areas of intense acetic acid uptake with distinct margins, usually close to or arising from the SCJ. If the TZ is fully visible, a woman with a positive VIA test result can be treated immediately with cryotherapy or thermal ablation, subject to certain requirements, in a single-visit screen-and-treat approach (see Section 5.1; WHO, 2013a, 2019), or may be referred for triage with colposcopy and treated in the conventional manner.

Table 4.2 Categories of results of visual inspection with acetic acid (VIA) examination

Test result	Clinical findings 1 minute after application of 3–5% acetic acid
Negative	No acetowhite lesions or faint acetowhite lesions due to squamous metaplasia or regenerating epithelium, cervicitis, inflammation; acetowhitening of polyps, Nabothian cysts; acetowhitening of the SCJ; satellite acetowhite lesions far away from the SCJ
Positive	Sharp, distinct, well-defined, dense (opaque/dull or oyster white) acetowhite area with or without raised margins touching the SCJ; leukoplakia and warts
Suspicious for cancer	Large chalky white acetowhite lesions obliterating the endocervical canal with irregular surface and raised and rolled-out margins; bleeding on touch; clinically visible ulcerative, cauliflower-like growth or ulcer

SCJ, squamocolumnar junction.
Table compiled by the Working Group.

VIA positivity rates vary considerably, partly because of the intrinsic subjectivity of the method (Almonte et al., 2015). The diagnostic accuracy of VIA has been shown to be variable and dependent on several factors, including the training and experience of the test provider, the adequacy of the light source, the concentration of acetic acid used, participant characteristics such as age (Castle et al., 2014; Raifu et al., 2017), the presence of infection with carcinogenic HPV types (Castle et al., 2014), and coexisting cervical inflammation (see Section 4.2.2).

(ii) Visual inspection with Lugol's iodine (VILI)

Lugol's iodine (5%) is relatively expensive. It can be prepared locally and should be discarded after 3–6 months. VILI may also be used as an adjunct to VIA and as an aid to precise treatment. Normal mature squamous epithelium takes up iodine and becomes a mahogany brown colour because of its high glycogen content. Dysplastic, metaplastic, and glandular epithelial tissues have minimal or no glycogen and do not take up iodine; they appear as well-defined, thick, mustard or saffron yellow areas. For women indicated for treatment, Lugol's iodine is valuable in demarcating the outer limit of the TZ, enabling the size of the TZ to be estimated so that the dimensions of the probe or the number of applications to be used can be calculated. Lugol's iodine is also a reasonably effective antiseptic agent (Sankaranarayanan & Wesley, 2003).

As observed for VIA, VILI has variable sensitivity, ranging from 50% (95% CI, 31–69%) to 100% (95% CI, 70–100%), and specificity, ranging from 69% (95% CI, 68–70%) to 97% (95% CI, 97–98%), for precancerous lesions (Catarino et al., 2018). In studies that evaluated VIA and VILI in head-to-head comparisons, the sensitivity of VILI for CIN2+ was higher than that of VIA (relative sensitivity, 1.11; 95% CI, 1.06–1.16), without significant loss in specificity (relative specificity, 0.98; 95% CI, 0.95–1.01). The higher sensitivity of VILI may be because the colour changes produced by the application of Lugol's iodine are more apparent visually than the whitening observed after the application of acetic acid.

(b) Strengths and limitations

The strengths and limitations of cervical screening using VIA are summarized in Table 4.3. Naked-eye examination of the cervix with acetic acid and/or Lugol's iodine as a means of detecting cervical precancer arose because of the absence or suboptimal performance of the screening methods used in high-income countries (i.e. cytology followed by colposcopy) when used in LMICs. VIA and VILI have several advantages. Any type of health-care worker can perform the test, and the results are available

Table 4.3 Strengths and limitations of cervical screening using visual inspection with acetic acid (VIA)

Strengths	Limitations
Simple, affordable, safe, and easy to learn and practise clinical testing, which requires minimal infrastructure and no or minimal laboratory support	Provider-dependent test outcome
Acetic acid is widely available and affordable	Test accuracy, particularly sensitivity, is highly variable in different settings and is dependent on training, supervision, and regular quality assurance
Different categories of health-care providers can learn and perform VIA	No standardized training and quality assurance methods for ensuring provider competency
Rapid, real-time test with immediately available test results, which enables a single-visit screen-and-treat approach or immediate triage with colposcopy or colposcopy-directed biopsy	Less accurate in postmenopausal women, because the SCJ recedes into the endocervical canal with increasing age
Low start-up and sustaining costs, which may enable use of the VIA screen-and-treat approach in primary care services	Moderate to low specificity to distinguish CIN2+ leads to resources being spent on unnecessary treatment of women who are free of precancerous lesions in a single-visit approach; leads to unnecessary investigations, such as colposcopy or biopsy, in settings where triage in VIA-positive women is done. Variable sensitivity leads to some women with CIN2+ or CIN3+ being incorrectly classified as disease-free
Focused visualization of the cervix enables early diagnosis of preclinical, asymptomatic early cervical cancer	Health and cost implications of overtreatment because of low specificity and/or missed cases because of low sensitivity

CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; SCJ, squamocolumnar junction; VIA, visual inspection with acetic acid.
Table compiled by the Working Group.

immediately, which enables a screen-and-treat protocol. The tests are laboratory-independent and inexpensive. Finally, a screening programme established using naked-eye examination will familiarize women and health-care providers with the concept of cancer prevention. VIA was formally endorsed by WHO in 2013 as a legitimate means of screening, particularly as part of a screen-and-treat approach in LMICs (WHO, 2013a). The application of 3% or 5% acetic acid is also used in some regions to determine eligibility for ablative treatment in women with a positive HPV test result, and also to determine the site, size, and type of the TZ.

The primary problem with naked-eye techniques is that they are highly subjective and consequently have variable sensitivity and specificity to detect precancer. Quality control and quality assurance for visual screening are important to maintain uniform and reproducible criteria for test positivity, and to ensure that the provider accurately differentiates between true-positive and true-negative cases (WHO, 2013b). Ensuring adequate training, supervision, and continuing quality assurance can be challenging in practice. Furthermore, visual examinations are an assessment of the ectocervical epithelium and cannot detect either glandular disease or endocervical squamous disease. In perimenopausal and postmenopausal women, the SCJ recedes into the endocervical canal and thus cannot be adequately observed with naked-eye examination. Even a proportion of women of reproductive age have a TZ of type 2 or 3 (see Fig. 1.18).

Table 4.4 Pooled sensitivity and specificity of visual inspection with acetic acid (VIA) to detect CIN2+ lesions

Reference	Study population Reference standard	Pooled sensitivity (%) (95% CI)	Pooled specificity (%) (95% CI)
Arbyn et al. (2008)	58 679 women from 11 studies Colposcopy with or without biopsy	79 (73–85)	85 (81–89)
Zhao et al. (2010)	28 848 women from 17 studies Four-quadrant biopsies	48 (42–54)	90 (87–94)
Chen et al. (2012)	99 972 women from 22 studies Colposcopy with or without biopsy	77 (75–78)	87 (87–88)
Bobdey et al. (2015)	57 225 women from 11 studies Colposcopy with or without biopsy	69 (32–100)	84 (53–91)
Fokom-Domgue et al. (2015)	61 381 women from 15 studies Colposcopy with or without biopsy	82 (76–87)	87 (78–93)
Adsul et al. (2017)	313 553 women from 20 studies Colposcopy with or without biopsy	17–83[a]	82–97[a]
Catarino et al. (2018)	101 273 women from 23 studies Colposcopy followed by colposcopy-directed biopsy or excision biopsy	78 (73–83)	88 (85–91)

CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse.
[a] Range in included studies; pooled estimates are not presented.

(c) Quality assurance for VIA

Quality assurance includes (but is not limited to) competency-based training of VIA providers, supervision, periodic refresher training, evaluation of current programme activities and long-term impact, a mechanism for constructive feedback from women and health-care providers, and an effective information system (see also Section 2.3). Training requirements for VIA providers are highly variable; WHO recommends a 10-day training (WHO, 2017), but in different programmatic settings the duration of training varies between 5 days and a few months (Blumenthal et al., 2005). Training is mainly non-standardized and is one of the weakest components of VIA screening initiatives. Some training manuals are available, which have been adapted by many countries (Sankaranarayanan & Wesley, 2003; WHO, 2013a, 2017), and a guide for quality control and quality assurance for VIA-based programmes has been published by WHO (WHO, 2013b).

4.2.2 Beneficial effects of screening using VIA

(a) Accuracy of VIA screening

VIA has been evaluated for its accuracy to detect CIN2+ lesions in cross-sectional studies in various settings in Africa, Asia, and Latin America. In most of these studies, the diagnostic reference standard used to establish the final diagnosis was colposcopy plus colposcopy-directed biopsy (Table 4.4), although some studies in China used four-quadrant biopsies to establish the final diagnosis (Belinson et al., 2001; Zhao et al., 2010, 2020; Holt et al., 2017). In studies that relied on colposcopy as the reference standard, no biopsies were directed when no colposcopic abnormalities were detected; directed biopsies were reserved for women with colposcopic abnormalities. In some studies the reference standard was used for all cases, thereby eliminating verification bias to a large extent, whereas in other studies the reference standard was used for all screen-positive women plus a proportion of screen-negative women. When the colposcopic

impression did not suggest precancer, no biopsy was taken, and this outcome was accepted as absence of precancer. [Given that standard colposcopy can miss up to 40.0% of prevalent precancers (Wentzensen et al., 2015), and given the inherent verification bias in studies and the close correlation of colposcopy with visual screening approaches, the reported sensitivity estimates of VIA are likely to be inflated.]

There is wide variation in VIA positivity rates across studies, from 1% to 36%. This indicates that VIA performance is subjective and depends on the study and the provider; there is little reproducibility, and provider training and thresholds used for test positivity vary (Jeronimo et al., 2014; Shastri et al., 2014; Huchko et al., 2015a, b; Poli et al., 2015).

In meta-analyses, the pooled sensitivity of VIA to detect CIN2+ lesions ranged from 48% to 83%, and the pooled specificity varied from 84% to 97% (Table 4.4). The sensitivity of VIA declines substantially in postmenopausal women. In a pooled analysis of 17 population-based studies in postmenopausal women, the sensitivity of VIA to detect CIN2+ lesions was 31.0% (95% CI, 21.8–41.4%) and the specificity was 94.6% (93.7–95.4%) (Holt et al., 2017). [The interpretation of VIA in perimenopausal and postmenopausal women is challenging, because the epithelium is pale, degenerated, and brittle and it bleeds on touch, and the TZ is partially visible or not visible. Given the methodological limitations, estimates of the absolute accuracy of VIA should be interpreted with caution.] It has been shown that low-level magnification does not improve the performance of naked-eye VIA (Basu et al., 2003; Sankaranarayanan et al., 2004; Shastri et al., 2005; Chen et al., 2012; Bobdey et al., 2015). Variation in test positivity is partly responsible for the varying accuracy of VIA in detecting high-grade lesions; the quality of the diagnostic reference standard used in different settings, which is also highly variable, is another factor that determines the variability of accuracy estimates (Sankaranarayanan et al., 2012).

HIV-positive women have a higher prevalence of HPV infection and a higher incidence of cervical cancer compared with HIV-negative women, partly because of the modifying effect of HIV on HPV pathogenesis (see Section 5.2.1). The screening methods used for HIV-seropositive women are the same as those used for HIV-negative women, with varying clinical performance and accuracy. In HIV-positive women, the use of VIA to detect CIN2+ had a sensitivity of 48.0–80.0% and a specificity of 65.0–92.0% (Ghebre et al., 2017; Mapanga et al., 2018). Visual screening tests might be expected to perform better in HIV-positive women than in the general population, because of the higher prevalence of high-grade lesions and the possibility of large lesions in HIV-positive women (Sahasrabuddhe et al., 2012; Joshi et al., 2013), although a high prevalence of HPV infection and other infections as well as inflammation may adversely affect the specificity of VIA.

(b) Cervical cancer incidence and mortality

VIA screening has been evaluated for its effect on cervical cancer incidence and/or mortality compared with control populations receiving usual care (very low prevalence of screening) in three large cluster-randomized trials in India (Sankaranarayanan et al., 2007, 2009; Shastri et al., 2014). The cervical cancer incidence rates, the detection rates of CIN2+ lesions, and the cervical cancer mortality rates in the VIA and control groups are given in Table 4.5. VIA positivity rates ranged from 2% (Shastri et al., 2014) to 13.9% (Sankaranarayanan et al., 2009), which indicates the subjective nature of VIA interpretation, differences in training and quality assurance, and possibly different thresholds used for VIA positivity.

In the study in Dindigul District, India, the intervention was a single round of VIA by trained nurses (Sankaranarayanan et al., 2007).

Table 4.5 Detection rates of CIN2/CIN3 lesions and cervical cancer incidence and mortality rates in randomized trials of screening with visual inspection with acetic acid (VIA)

Reference Study design	Cervical cancer incidence rate per 100 000 person-years		Detection rate of CIN2/CIN3 lesions per 1000 women invited		Screen-negative cervical cancer incidence rate per 100 000 person-years	Cervical cancer mortality rate per 100 000 person-years	
	VIA group	Control group	VIA group	Control group		VIA group	Control group
Sankaranarayanan et al. (2007) 49 311 women aged 30–59 yr in the VIA group and 30 958 women in the control group; single round of VIA screening by nurses	75.2[a]	99.1[a]	4.84	NA	NA	39.6[a]	56.7[a]
Sankaranarayanan et al. (2009) 34 074 women aged 30–59 yr in the VIA group and 31 488 women in the control group; single round of VIA screening by trained auxiliary-nurse midwives	58.7[a]	47.6[a]	5.72	0.48	16.0[b]	20.9[a]	25.8[a]
Shastri et al. (2014) 75 360 women aged 30–64 yr in the VIA group and 76 178 women in the control group; 4 rounds of VIA at 2-yr intervals by primary health workers	29.0[a]	29.4[a]	1.44	0.17	NA	14.4[a]	19.8[a]

CIN2, cervical intraepithelial neoplasia grade 2; CIN3, cervical intraepithelial neoplasia grade 3; NA, not available; VIA, visual inspection with acetic acid; yr, year or years.

[a] Standardized rate using world standard population.

[b] Invasive cervical cancer.

The study involved women aged 30–59 years, with 49 311 in the VIA group and 30 958 in the control group. Of the 3088 (9.9%) women with a positive test result on VIA, 3052 (98.9%) underwent colposcopy and 2539 (82.2%) had directed biopsy. Of the 1874 women with precancerous lesions, 72.0% received treatment. During 2000–2006, in the VIA group, for 274 430 person-years, 167 cervical cancer cases and 83 cervical cancer deaths were recorded, whereas in the control group, for 178 781 person-years, 158 cervical cancer cases and 92 cervical cancer deaths were recorded (incidence hazard ratio, 0.75; 95% CI, 0.55–0.95; mortality hazard ratio, 0.65; 95% CI, 0.47–0.89). The Dindigul District study was the first randomized trial of VIA screening to report a significant reduction in cervical cancer incidence and mortality after VIA screening.

In the study in Osmanabad District, India, a single round of VIA was administered by trained paramedical workers (Sankaranarayanan et al., 2009). The study involved women aged 30–59 years, with 34 074 in the VIA group and 31 488 in the control group. In the VIA group, the VIA positivity rate was 13.9%; this decreased from 17.8% in women aged 30–39 years to 6.4% in women aged 50–59 years. In the VIA group, 195 women with CIN2 and CIN3 lesions, 157 cervical cancers, and 56 cervical cancer deaths were recorded, whereas in the control group 15 women with CIN2 and CIN3 lesions, 118 cervical cancers, and 64 cervical cancer deaths were recorded (incidence hazard ratio, 1.30; 95% CI,

0.95–1.78; mortality hazard ratio, 0.86; 95% CI, 0.60–1.25) (Sankaranarayanan et al., 2009).

[The differing results for VIA screening in the two above-mentioned studies may be due to a lack of power to detect a significant reduction in mortality in the Osmanabad District study and the higher frequency of treatment of precancerous lesions in the Dindigul District study. In the Osmanabad District study, screening with HPV testing was associated with a significant reduction in advanced disease and mortality, indicating a better accuracy to detect precancerous lesions.]

The third trial, in Mumbai, evaluated four rounds of VIA screening provided by trained primary health workers every 2 years (Shastri et al., 2014). The VIA positivity rate varied from 1.3% to 2.5%. This study recruited 75 360 women aged 30–64 years from 10 clusters in the VIA group and 76 178 women from 10 comparable clusters in the control group. A significant 31% reduction in cervical cancer mortality (incidence rate ratio [IRR], 0.69; 95% CI, 0.54–0.88; $P = 0.003$) and a non-significant 7% reduction in all-cause mortality (mortality IRR, 0.93; 95% CI, 0.79–1.10; $P = 0.41$) was associated with VIA screening compared with the control group, but no reduction in the incidence of cervical cancer was observed (IRR, 0.97; 95% CI, 0.80–1.19; $P = 0.79$). [The low detection rate of high-grade lesions, possibly as a consequence of low VIA positivity rates (1.3–2.5%) in the four rounds of VIA screening, along with stage shift of invasive cancers, possibly led to the reduction in mortality only rather than reductions in both incidence and mortality in the Mumbai trial.

(c) Single-visit VIA screen-and-treat approach

In an RCT in women aged 35–65 years in South Africa, HPV DNA screen-and-treat (2163 women) and VIA screen-and-treat (2227 women) protocols were compared with a delayed-evaluation group (2165 women). At 6 months after randomization, the prevalence of CIN2+ lesions was significantly lower in the two screen-and-treat groups than in the delayed-evaluation group (Denny et al., 2005). In both screened groups, 22% of women underwent cryotherapy. At 6 months, CIN2+ lesions were detected in 2.23% (95% CI, 1.57–2.89%) of women in the VIA group compared with 3.55% (95% CI, 2.71–4.39%) of women in the delayed-evaluation group ($P = 0.02$); in the HPV DNA group, CIN2+ lesions were detected in 0.80% (95% CI, 0.40–1.20%) of women. At 12 months, the cumulative prevalence of CIN2+ lesions in a subset of women was 2.91% (95% CI, 2.12–3.69%) in the VIA group and 5.41% (95% CI, 4.32–6.50%) in the delayed-evaluation group; in the HPV DNA group, the cumulative prevalence of CIN2+ lesions was 1.42% (95% CI, 0.87–1.97%). There were no differences in HIV seroconversion rates 6 months after randomization; this was reassuring about possible virus transmission during screen-and-treat procedures, but the study was underpowered to detect small increases.

4.2.3 Harms of screening using VIA

Although VIA has been evaluated for its performance in cross-sectional studies in Africa, Asia, and Latin America and has been implemented opportunistically as a point-of-care screening approach or in programmes, there is very little systematic documentation of associated harms (Muwonge et al., 2010; Poli et al., 2015). Given the simplicity of VIA as a screening procedure, the innocuous nature of acetic acid, and the lack of documentation of serious adverse events in studies, VIA is assumed to be safe. A few studies have documented the rate of important potential harms, including adverse reproductive outcomes (from treatment) and complications that can be directly attributed to VIA, although the evidence is of low quality (Fokom-Domgue et al., 2014). Arguably the major risk of VIA as a screening test is that it will not always recognize

Box 4.1 Harms of visual screening

- Physical harms associated with true-positive test results (i.e. accurate screening, correct diagnosis and treatment):
  - pain and discomfort during screening and treatment
  - discharge, pain, bleeding, and infection risk after treatment
  - long-term treatment complications (premature labour, threatened miscarriage, and cervical stenosis)
- Psychological harms:
  - periprocedural anxiety
  - psychological stress and fear of pelvic examination, VIA screening, and downstream procedures of diagnosis, treatment, and follow-up care
- Harms associated with false-positive test results:
  - unnecessary investigations (if triage of women with a positive test result is done)
  - unnecessary biopsy
  - overtreatment (with attendant risk of short-term and long-term physical harms as detailed above)
  - costs of unnecessary medical care
- False reassurance and risk of future cervical neoplasia because of a false-negative test result
- Harms associated with overdiagnosis

an endocervical TZ and thus may falsely reassure a woman that she does not have precancer when in fact she does.

(a) Physical harms

There is very little documentation of either immediate physical harm (such as bleeding, pain and irritation due to insertion of the speculum, lower abdominal cramps, syncope, febrile illness, or allergic reactions) or late adverse events (such as delayed bleeding, cervicitis, cervical ulceration, pelvic inflammatory disease, pregnancy loss, preterm labour, or cervical stenosis) from examination with VIA.

Given the well-documented limitations in the accuracy of VIA, there are likely to be harms from overtreatment of women with false-positive test results (Parra et al., 2020), particularly in the screen-and-treat setting, as well as the potentially serious harm of a failure to detect a lesion that may develop into invasive cancer (false-negative test result). Potential harms of false-positive and false-negative test results are given in Box 4.1. False-positive test results lead to unnecessary investigations and costs of unnecessary medical care (in settings using triage with colposcopy of women with a positive test result), unnecessary biopsy, and harms associated with treatment, such as excessive discharge, risks of bleeding, infection and pelvic inflammatory disease, and long-term sequelae such as premature labour, threatened miscarriage, and cervical stenosis. [Variations in the accuracy of visual screening are caused by variations in the performance of VIA providers rather than underlying variations in the prevalence of disease; this indicates

that harms associated with VIA can be reduced if providers are well trained in the procedure (Raifu et al., 2017).]

(b) Psychological harms

Psychological harms include anxiety and fear caused by the procedure itself and by a positive test result, and the stress associated with making the decision to accept screen-and-treat in the same session (in a single-visit approach) and to give consent for eligibility determination and treatment procedures. Women undergoing pelvic examination can experience anxiety, fear, and embarrassment, and the associated stress can lead to exacerbation of procedure-related discomfort, which may discourage women from undergoing the procedure and may induce low patient compliance (Galaal et al., 2011; O'Connor et al., 2016a, b; Vorsters et al., 2017). In one study in Cameroon, enabling women to watch the VIA procedure on a digital screen in real time improved their emotional state but did not reduce periprocedural anxiety as measured by the Spielberger State-Trait Anxiety Inventory (STAI) score (Camail et al., 2019).

References

Adsul P, Manjunath N, Srinivas V, Arun A, Madhivanan P (2017). Implementing community-based cervical cancer screening programs using visual inspection with acetic acid in India: a systematic review. *Cancer Epidemiol.* 49:161–74. doi:10.1016/j.canep.2017.06.008 PMID:28704717

Almonte M, Ferreccio C, Luciani S, Gonzales M, Delgado JM, Santos C, et al. (2015). Visual inspection after acetic acid (VIA) is highly heterogeneous in primary cervical screening in Amazonian Peru. *PLoS One.* 10(1):e0115355. doi:10.1371/journal.pone.0115355 PMID:25635965

Arbyn M, Sankaranarayanan R, Muwonge R, Keita N, Dolo A, Mbalawa CG, et al. (2008). Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *Int J Cancer.* 123(1):153–60. doi:10.1002/ijc.23489 PMID:18404671

Basu PS, Sankaranarayanan R, Mandal R, Roy C, Das P, Choudhury D, et al.; Calcutta Cervical Cancer Early Detection Group (2003). Visual inspection with acetic acid and cytology in the early detection of cervical neoplasia in Kolkata, India. *Int J Gynecol Cancer.* 13(5):626–32. doi:10.1136/ijgc-00009577-200309000-00009 PMID:14675346

Belinson J, Qiao YL, Pretorius R, Zhang WH, Elson P, Li L, et al. (2001). Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. [Erratum in: Gynecol Oncol. 2002 84(2):355] *Gynecol Oncol.* 83(2):439–44. doi:10.1006/gyno.2001.6370 PMID:11606114

Blumenthal PD, Lauterbach M, Sellors JW, Sankaranarayanan R (2005). Training for cervical cancer prevention programs in low-resource settings: focus on visual inspection with acetic acid and cryotherapy. *Int J Gynaecol Obstet.* 89(Suppl 2):S30–7. doi:10.1016/j.ijgo.2005.01.012 PMID:15823264

Bobdey S, Balasubramanium G, Kumar A, Jain A (2015). Cancer screening: should cancer screening be essential component of primary health care in developing countries? *Int J Prev Med.* 6:56. PMID:26236443

Camail R, Kenfack B, Tran PL, Viviano M, Tebeu PM, Temogne L, et al. (2019). Benefit of watching a live visual inspection of the cervix with acetic acid and Lugol iodine on women's anxiety: randomized controlled trial of an educational intervention conducted in a low-resource setting. *JMIR Cancer.* 5(1):e9798. doi:10.2196/cancer.9798 PMID:31094335

Castle PE, Qiao YL, Zhao FH, Chen W, Valdez M, Zhang X, et al. (2014). Clinical determinants of a positive visual inspection after treatment with acetic acid for cervical cancer screening. *BJOG.* 121(6):739–46. doi:10.1111/1471-0528.12646 PMID:24575872

Catarino R, Schäfer S, Vassilakos P, Petignat P, Arbyn M (2018). Accuracy of combinations of visual inspection using acetic acid or Lugol iodine to detect cervical precancer: a meta-analysis. *BJOG.* 125(5):545–53. doi:10.1111/1471-0528.14783 PMID:28603909

Chen C, Yang Z, Li Z, Li L (2012). Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis. *Int J Gynecol Cancer.* 22(6):908–21. doi:10.1097/IGC.0b013e318256e5e4 PMID:22672987

Denny L, Kuhn L, De Souza M, Pollack AE, Dupree W, Wright TC Jr (2005). Screen-and-treat approaches for cervical cancer prevention in low-resource settings: a randomized controlled trial. *JAMA.* 294(17):2173–81. doi:10.1001/jama.294.17.2173 PMID:16264158

Fokom-Domgue J, Combescure C, Fokom-Defo V, Tebeu PM, Vassilakos P, Kengne AP, et al. (2015). Performance of alternative strategies for primary cervical cancer screening in sub-Saharan Africa: systematic review and meta-analysis of diagnostic

test accuracy studies. *BMJ*. 351:h3084. doi:10.1136/bmj.h3084 PMID:26142020

Fokom-Domgue J, Vassilakos P, Petignat P (2014). Is screen-and-treat approach suited for screening and management of precancerous cervical lesions in sub-Saharan Africa? *Prev Med*. 65:138–40. doi:10.1016/j.ypmed.2014.05.014 PMID:24879892

Galaal K, Bryant A, Deane KHO, Al-Khaduri M, Lopes AD (2011). Interventions for reducing anxiety in women undergoing colposcopy. *Cochrane Database Syst Rev*. (12):CD006013. doi:10.1002/14651858.CD006013.pub3 PMID:22161395

Ghebre RG, Grover S, Xu MJ, Chuang LT, Simonds H (2017). Cervical cancer control in HIV-infected women: past, present and future. *Gynecol Oncol Rep*. 21:101–8. doi:10.1016/j.gore.2017.07.009 PMID:28819634

Goldstein A, Goldstein LS, Lipson R, Bedell S, Wang J, Stamper SA, et al. (2020). Assessing the feasibility of a rapid, high-volume cervical cancer screening programme using HPV self-sampling and digital colposcopy in rural regions of Yunnan, China. *BMJ Open*. 10(3):e035153. doi:10.1136/bmjopen-2019-035153 PMID:32234744

Holt HK, Zhang L, Zhao FH, Hu SY, Zhao XL, Zhang X, et al. (2017). Evaluation of multiple primary and combination screening strategies in postmenopausal women for detection of cervical cancer in China. *Int J Cancer*. 140(3):544–54. doi:10.1002/ijc.30468 PMID:27727464

Huchko MJ, Sneden J, Sawaya G, Smith-McCune K, Maloba M, Abdulrahim N, et al. (2015a). Accuracy of visual inspection with acetic acid to detect cervical cancer precursors among HIV-infected women in Kenya. *Int J Cancer*. 136(2):392–8. doi:10.1002/ijc.28996 PMID:24889387

Huchko MJ, Sneden J, Zakaras JM, Smith-McCune K, Sawaya G, Maloba M, et al. (2015b). A randomized trial comparing the diagnostic accuracy of visual inspection with acetic acid to visual inspection with Lugol's iodine for cervical cancer screening in HIV-infected women. *PLoS One*. 10(4):e0118568. doi:10.1371/journal.pone.0118568 PMID:25849627

Jeronimo J, Bansil P, Lim J, Peck R, Paul P, Amador JJ, et al.; START-UP Study Group (2014). A multicountry evaluation of careHPV testing, visual inspection with acetic acid, and Papanicolaou testing for the detection of cervical cancer. *Int J Gynecol Cancer*. 24(3):576–85. doi:10.1097/IGC.0000000000000084 PMID:24557438

Joshi S, Sankaranarayanan R, Muwonge R, Kulkarni V, Somanathan T, Divate U (2013). Screening of cervical neoplasia in HIV-infected women in India. *AIDS*. 27(4):607–15. doi:10.1097/QAD.0b013e32835b1041 PMID:23079814

Mapanga W, Girdler-Brown B, Feresu SA, Chipato T, Singh E (2018). Prevention of cervical cancer in HIV-seropositive women from developing countries through cervical cancer screening: a systematic review. *Syst Rev*. 7(1):198. doi:10.1186/s13643-018-0874-7 PMID:30447695

Muwonge R, Manuel MG, Filipe AP, Dumas JB, Frank MR, Sankaranarayanan R (2010). Visual screening for early detection of cervical neoplasia in Angola. *Int J Gynaecol Obstet*. 111(1):68–72. doi:10.1016/j.ijgo.2010.04.024 PMID:20570259

O'Connor M, Gallagher P, Waller J, Martin CM, O'Leary JJ, Sharp L; Irish Cervical Screening Research Consortium (CERVIVA) (2016b). Adverse psychological outcomes following colposcopy and related procedures: a systematic review. *BJOG*. 123(1):24–38. doi:10.1111/1471-0528.13462 PMID:26099164

O'Connor M, O'Leary E, Waller J, Gallagher P, D'arcy T, Flannelly G, et al.; Irish Cervical Screening Research Consortium (CERVIVA) (2016a). Trends in, and predictors of, anxiety and specific worries following colposcopy: a 12-month longitudinal study. *Psychooncology*. 25(5):597–604. doi:10.1002/pon.3980 PMID:26392040

Parham GP, Mwanahamuntu MH, Kapambwe S, Muwonge R, Bateman AC, Blevins M, et al. (2015). Population-level scale-up of cervical cancer prevention services in a low-resource setting: development, implementation, and evaluation of the cervical cancer prevention program in Zambia. *PLoS One*. 10(4):e0122169. doi:10.1371/journal.pone.0122169 PMID:25885821

Parra SG, López-Orellana LM, Molina Duque AR, Carns JL, Schwarz RA, Smith CA, et al. (2020). Cervical cancer prevention in El Salvador: a prospective evaluation of screening and triage strategies incorporating high-resolution microendoscopy to detect cervical precancer. *Int J Cancer*. doi:10.1002/ijc.33454 PMID:33368249

Poli UR, Bidinger PD, Gowrishankar S (2015). Visual inspection with acetic acid (VIA) screening program: 7 years experience in early detection of cervical cancer and pre-cancers in rural south India. *Indian J Community Med*. 40(3):203–7. doi:10.4103/0970-0218.158873 PMID:26170547

Raifu AO, El-Zein M, Sangwa-Lugoma G, Ramanakumar A, Walter SD, Franco EL; Congo Screening Study Group (2017). Determinants of cervical cancer screening accuracy for visual inspection with acetic acid (VIA) and Lugol's iodine (VILI) performed by nurse and physician. *PLoS One*. 12(1):e0170631. doi:10.1371/journal.pone.0170631 PMID:28107486

Sahasrabuddhe VV, Bhosale RA, Kavatkar AN, Nagwanshi CA, Joshi SN, Jenkins CA, et al. (2012). Comparison of visual inspection with acetic acid and cervical cytology to detect high-grade cervical neoplasia among HIV-infected women in India. *Int J Cancer*. 130(1):234–40. doi:10.1002/ijc.25971 PMID:21387289

Sankaranarayanan R, Esmy PO, Rajkumar R, Muwonge R, Swaminathan R, Shanthakumari S, et al. (2007). Effect of visual screening on cervical cancer incidence and mortality in Tamil Nadu, India: a cluster-randomised

trial. *Lancet*. 370(9585):398–406. doi:10.1016/S0140-6736(07)61195-7 PMID:17679017

Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. (2009). HPV screening for cervical cancer in rural India. *N Engl J Med*. 360(14):1385–94. doi:10.1056/NEJMoa0808516 PMID:19339719

Sankaranarayanan R, Nessa A, Esmy PO, Dangou JM (2012). Visual inspection methods for cervical cancer prevention. *Best Pract Res Clin Obstet Gynaecol*. 26(2):221–32. doi:10.1016/j.bpobgyn.2011.08.003 PMID:22075441

Sankaranarayanan R, Shastri SS, Basu P, Mahé C, Mandal R, Amin G, et al. (2004). The role of low-level magnification in visual inspection with acetic acid for the early detection of cervical neoplasia. *Cancer Detect Prev*. 28(5):345–51. doi:10.1016/j.cdp.2004.04.004 PMID:15542259

Sankaranarayanan R, Wesley R, Somanathan T, Dhakad N, Shyamalakumary B, Amma NS, et al. (1998). Visual inspection of the uterine cervix after the application of acetic acid in the detection of cervical carcinoma and its precursors. *Cancer*. 83(10):2150–6. doi:10.1002/(SICI)1097-0142(19981115)83:10<2150::AID-CNCR13>3.0.CO;2-0 PMID:9827719

Sankaranarayanan R, Wesley RS, editors (2003). A practical manual on visual screening for cervical neoplasia (IARC Technical Publications, No. 41). Lyon, France: International Agency for Research on Cancer. Available from: https://publications.iarc.fr/364.

Shastri SS, Dinshaw K, Amin G, Goswami S, Patil S, Chinoy R, et al. (2005). Concurrent evaluation of visual, cytological and HPV testing as screening methods for the early detection of cervical neoplasia in Mumbai, India. *Bull World Health Organ*. 83(3):186–94. PMID:15798842

Shastri SS, Mittra I, Mishra GA, Gupta S, Dikshit R, Singh S, et al. (2014). Effect of VIA screening by primary health workers: randomized controlled study in Mumbai, India. *J Natl Cancer Inst*. 106(3):dju009. doi:10.1093/jnci/dju009 PMID:24563518

Vorsters A, Arbyn M, Baay M, Bosch X, de Sanjosé S, Hanley S, et al. (2017). Overcoming barriers in HPV vaccination and screening programs. *Papillomavirus Res*. 4:45–53. doi:10.1016/j.pvr.2017.07.001 PMID:29179869

Wentzensen N, Walker JL, Gold MA, Smith KM, Zuna RE, Mathews C, et al. (2015). Multiple biopsies and detection of cervical cancer precursors at colposcopy. *J Clin Oncol*. 33(1):83–9. doi:10.1200/JCO.2014.55.9948 PMID:25422481

WHO (2013a). WHO guidelines for screening and treatment of precancerous lesions for cervical cancer prevention. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/94830https://www.who.int/reproductivehealth/publications/cancers/screening_and_treatment_of_precancerous_lesions/en/.

WHO (2013b). Monitoring national cervical cancer prevention and control programmes: quality control and quality assurance for visual inspection with acetic acid (VIA)-based programmes. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/79316.

WHO (2014). Comprehensive cervical cancer control: a guide to essential practice. 2nd ed. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/144785.

WHO (2017). Training of health staff in VIA, HPV detection test and cryotherapy – trainees' handbook. New Delhi, India: World Health Organization Regional Office for South-East Asia. Available from: https://apps.who.int/iris/handle/10665/259001.

WHO (2019). WHO guidelines for the use of thermal ablation for cervical pre-cancer lesions. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/329299.

Xue Z, Novetsky AP, Einstein MH, Marcus JZ, Befano B, Guo P, et al. (2020). A demonstration of automated visual evaluation of cervical images taken with a smartphone camera. *Int J Cancer*. 147(9):2416–23. doi:10.1002/ijc.33029 PMID:32356305

Zhao FH, Lin MJ, Chen F, Hu SY, Zhang R, Belinson JL, et al.; Cervical Cancer Screening Group in China (2010). 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*. 11(12):1160–71. doi:10.1016/S1470-2045(10)70256-4 PMID:21075054

Zhao XL, Xu XQ, Duan XZ, Rezhake R, Hu SY, Wang Y, et al. (2020). Comparative performance evaluation of different HPV tests and triaging strategies using self-samples and feasibility assessment of thermal ablation in 'colposcopy and treat' approach: a population-based study in rural China. *Int J Cancer*. 147(5):1275–85. doi:10.1002/ijc.32881 PMID:31970767

4.3 Cytological methods

4.3.1 Technical descriptions

Cytology is an established method of primary screening that is used to identify preclinical lesions and prevent the development of invasive cancer (Morrison, 1992). The technique of cervical cytology was developed by Papanicolaou and Babeş in the 1920s and later improved by Papanicolaou (Swailes et al., 2019). In the 1960s, cervical cytology was adopted for cervical cancer screening and was introduced in some high-income countries. Since then, the primary aim of the Pap test has shifted from the detection of invasive cancer to the identification of precancerous lesions. The main method used in primary screening has changed from cytology to HPV testing, particularly in Australia, some European countries, and the USA (Cuschieri et al., 2018) (see Section 2.2). However, in some countries, cytology still has a significant role in primary screening and triage. To reduce unnecessary colposcopy, a triage step has been introduced after the detection of low-grade abnormalities (see Section 4.4.7). Although cytology is used for this purpose, HPV testing, p16/Ki-67 dual staining, and some molecular biomarkers have been adopted as alternative methods.

(a) Conventional cytology

The conventional cytology technique involves collecting exfoliated cells from the TZ and endocervical canal. The precursors of cervical squamous cell carcinoma (SCC) occur mainly in the transformation zone (Burghardt, 1970). Thus, endocervical and/or metaplastic cells from the transformation zone are necessary for the adequacy of the sample (Arbyn et al., 2008a). However, the absence of endocervical cells is not necessarily associated with a high risk of future cervical neoplasia (Mitchell, 2001; McCredie et al., 2008; Sultana et al., 2014).

The quality of the smear is an essential component of the cytological interpretation. If too few cells are taken, the sample will not be representative of cells from the cervix (National Institute for Health and Clinical Excellence, 2003) and will be classified as unsatisfactory, because it cannot be interpreted. Unsatisfactory samples prevent the microscopic evaluation. A cervical sample is usually taken by a health service provider, such as a gynaecologist, general physician, midwife, or trained nurse (McDonald et al., 2001; Ideström et al., 2007; Yabroff et al., 2009; Cooper & Saraiya, 2014). Training of health providers in smear collection to ensure that samples are of adequate quantity and quality plays a critical role in quality assurance (see Section 4.3.1f).

Ideally, cytological examinations should be performed about 2 weeks after the first day of the previous menstrual period (IARC, 2005; Arbyn et al., 2008a). Sexual intercourse within 24 hours and use of intravaginal estrogen products should be avoided before cytological examinations. After childbirth, it is difficult to take adequate cervical samples for interpretation until 8 weeks postpartum.

The use of an appropriate collection device is essential in helping to reduce the proportion of unsatisfactory smears. Various instruments are used for taking smears, including cotton swabs, wooden spatulas, plastic spatulas, cytobrushes, and cervical brooms (Cervex-Brush). A study in Japan reported that before the introduction of the Bethesda system, more than 10% of smears collected using cotton swabs were reported as unsatisfactory (Hosono et al., 2018). Martin-Hirsch et al. (2000) compared collection devices for obtaining cytological samples in a systematic review of randomized and non-randomized comparative studies. The cervical broom is a commonly used device, and it was found that smears taken with it are adequate and comparable to those taken with a spatula (Peto odds ratio, 1.08; 95% CI, 0.97–1.21). However, a spatula

with an attached cytobrush performed better than the cervical broom alone (Peto odds ratio, 1.52; 95% CI, 1.15–2.01).

Cells collected for microscopic examination are applied to a glass slide for conventional cytology and commonly fixed using 95% ethyl alcohol covering the whole cellular area of the slide (Arbyn et al., 2008a). Cell fixation is performed within a few seconds of specimen collection to prevent air-drying, which obscures cellular detail and hinders interpretation (Somrak et al., 1990). The conventional Pap test technique may sometimes result in unsatisfactory smears, which are difficult to interpret because of uneven cell distribution, overlapping cells, blood, or inflammation (Taylor et al., 2006; Ronco et al., 2006, 2007).

(b) Liquid-based cytology

Liquid-based cytology (LBC) is a more recent technique for transferring the cellular material to the microscope slide (Arbyn et al., 2008a). The brush with the sample is rinsed into a vial with preservative fluid and then transported to the laboratory (Siebers et al., 2009). This results in cells that better represent the sample being transferred to the glass slide when compared with conventional cytology (Payne et al., 2000). An LBC preparation more consistently results in a monolayer and reduces the proportion of unsatisfactory slides by avoiding transfer of blood and mucus. The subsequent process for staining and microscopic assessment of a slide is similar to that used in conventional cytology. However, LBC enables improved fixation, which leads to more consistent staining; this contributes to improved quality and readability. Training in the preparation technique and in the interpretation of LBC-specific slides is required for medical staff and cytologists (Payne et al., 2000). A major advantage of LBC over conventional cytology is that residual cell material can be used for additional testing, including testing for HPV and molecular biomarkers. A disadvantage is the need for specific equipment for LBC and the substantial increase in unit costs (Payne et al., 2000; Taylor et al., 2006; Arbyn et al., 2008a). Several materials for LBC are available as commercial systems, for example the ThinPrep Imaging System and the BD FocalPoint GS Imaging System (SurePath).

LBC has been reported to reduce the rate of unsatisfactory samples in some population-based programmes. In a population-based cervical cancer screening programme in the Netherlands, unsatisfactory rates were reported to be 0.89% for conventional cytology and 0.13% for LBC (Beerman et al., 2009). In England, a pilot study reported that the rate of unsatisfactory samples decreased from 9.1% with Pap smears to 1.6% with LBC; in Scotland, the decrease was from 13.6% to 1.9% (National Institute for Health and Clinical Excellence, 2003; Williams, 2006). However, recent reports from Asian countries have suggested that there was no significant difference between LBC and conventional cytology in the rate of unsatisfactory smears (Kituncharoen et al., 2015; Hosono et al., 2018). A low rate of unsatisfactory smears in conventional cytology may reflect a good quality assurance system (Schneider et al., 2000; Petry et al., 2003; Klug et al., 2013). In 9 of 11 RCTs, the rate of unsatisfactory cytology was halved using LBC compared with conventional cytology (see Section 4.3.3, Table 4.15).

When LBC is used, the samples taken can be used for additional investigations, such as HPV testing, without needing to recall the woman (Cox, 2009; Albrow et al., 2012). LBC has been used with HPV testing as a primary screening method or for triage of HPV-positive women. When co-testing was used, the detection rate of CIN2+ increased, but rates of referral for colposcopy doubled compared with LBC alone (Kitchener et al., 2009). When LBC was used to triage HPV-positive women, the detection rate was increased and there was also an increase in the rate of colposcopy referrals compared

with LBC screening followed by HPV triage of abnormal LBC (Ogilvie et al., 2017).

A major problem with LBC is the high cost of the equipment and consumables required for the established commercial LBC methods; this is a considerable barrier to its use in resource-constrained settings (Arbyn et al., 2008a; Gupta et al., 2017; Pankaj et al., 2018).

A manual method for LBC was developed by Maksem et al. (2001). Nandini et al. (2012) reported that the concordance between manual LBC and histopathology was improved compared with CC. Because manual LBC is less expensive than commercial LBC systems, it might be a good alternative in low-resource settings.

(c) Computer-assisted cytology

Computer-assisted screening systems for both conventional cytology and LBC have been available since the early 2000s; these enable rapid interpretation of slides, which means that fewer professionals are needed (Thrall, 2019). In particular, some of these systems were developed to rapidly identify slides with normal cytology results that do not require further manual review.

The sensitivity and specificity of the PAPNET system, the first computer-assisted system for conventional cytology, was reported to be equal to that of conventional cervical screening (Doornewaard et al., 1999; Duggan, 2000). In population-based screening in the Netherlands, Kok & Boon (1996) reported that the diagnosis of HSIL and invasive cancer was higher for PAPNET than for conventional cytology. A study in Finland was the first RCT to evaluate the efficacy of automated screening using PAPNET (Nieminen et al., 2003, 2007; Anttila et al., 2011). More cases of LSIL were detected by screening with computer-assisted than with conventional cytology (RR, 1.08; 95% CI, 1.01–1.15), and significantly more cases of CIN1+ were detected with computer-assisted cytology (RR, 1.11; 95% CI, 1.02–1.21) (Nieminen et al., 2007). However, after 6.3 years of follow-up, no difference was found in the risk of cervical cancer (RR, 1.00; 95% CI, 0.76–1.29) or of death from cervical cancer (RR, 1.11; 95% CI, 0.62–1.92) (Anttila et al., 2011).

For two more recently developed systems, ThinPrep and FocalPoint/SurePath, sensitivity and specificity were assessed by comparing the results with manual diagnosis by experts of the same slides (Biscotti et al., 2005; Wilbur et al., 2009). The sensitivities and specificities were nearly equivalent even when the test threshold was changed (Table 4.6).

A study in Australia evaluated the detection and unsatisfactory rate of the ThinPrep imager on the basis of 55 164 split-sample pairs (Davey et al., 2007). There were fewer unsatisfactory slides with the ThinPrep imager than with conventional cytology. LBC with the ThinPrep imager detected 1.3 more cases of high-grade lesions per 1000 women screened than conventional cytology.

The Manual Assessment Versus Automated Reading In Cytology (MAVARIC) trial was conducted to compare two automated systems (ThinPrep and FocalPoint/SurePath) with manual screening for the introduction of national programmes in England (Kitchener et al., 2011). The relative sensitivities of automated systems for CIN2+ compared with manual screening were nearly equal (ThinPrep relative sensitivity, 0.92; 95% CI, 0.87–0.98; FocalPoint relative sensitivity, 0.90; 95% CI, 0.85–0.96).

In an RCT in Germany, manual and automated LBC systems were compared (Klug et al., 2013). The relative sensitivity with LSIL as the threshold was 3.17 (95% CI, 1.94–5.19) for CIN2+ detection and 3.38 (95% CI, 3.38–6.21) for CIN3+ detection. Although the automated LBC system detected more CIN, the PPVs were equivalent. The relative PPV was 1.07 (95% CI, 0.75–1.53) for CIN2+ detection and 1.09 (95% CI, 0.66–1.80) for CIN3+ detection. In Denmark, Rebolj et al. (2015) assessed CIN detection rates and false-positive rates of LBC and computer-assisted reading based on routine screening data in a real-world

Table 4.6 Systematic reviews of studies of test performance of manual diagnosis compared with automated screening[a]

Test threshold	Manual		Automated (ThinPrep)		Reference
	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	
ASC-US	75.6 (72.2–78.8)	97.6 (97.2–97.9)	82.0 (78.8–84.8)	97.8 (97.2–97.9)	Biscotti et al. (2005)
LSIL	79.7 (75.3–83.7)	99.0 (98.8–99.2)	79.2 (74.7–83.2)	99.1 (98.9–99.3)	Biscotti et al. (2005)
HSIL	74.1 (66.0–81.2)	99.4 (99.2–99.6)	79.9 (72.2–86.2)	99.6 (99.5–99.7)	Biscotti et al. (2005)
	Manual		Automated (FocalPoint)		
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	
ASC-US	82.6	82.7	81.1	84.5	Wilbur et al. (2009)
LSIL	76.4	90.6	86.1	88.7	Wilbur et al. (2009)
HSIL	65.7	97.7	85.3	95.1	Wilbur et al. (2009)

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.

[a] Reference standards in both studies were defined as the diagnosis of cytology carried out by experts in each study.

setting. For women aged 23–29 years with an atypical squamous cells of undetermined significance (ASC-US) threshold, the FocalPoint/SurePath system significantly increased the detection of CIN3+ (relative sensitivity, 1.85; 95% CI, 1.55–2.21) compared with manually read conventional cytology, but the increase was not significant using ThinPrep (relative sensitivity, 1.11; 95% CI, 0.88–1.39). The detection rate and false-positive rate of automated LBC depended upon brand and age group.

(d) The Bethesda system

The Bethesda system (TBS) is widely used for reporting cervical cytological diagnoses, but the Pap and WHO systems are also used in some areas. The relationship between the systems currently in use is shown in Fig. 1.17 (see also Section 1.2.3). In TBS 2001, the results of smears are assessed for specimen adequacy and divided into three categories: negative for intraepithelial lesion or malignancy (NILM), epithelial cell abnormalities (with either squamous cells or glandular cells), and others. Squamous cell abnormalities are classified as follows: ASC-US; atypical squamous cells, cannot exclude HSIL (ASC-H); LSIL; HSIL; and SCC. Of women with atypical squamous cells (ASC), 10–20% have underlying CIN2 or CIN3 and 0.1% have invasive cancer (Solomon et al., 2001). Specific glandular cell abnormalities are classified as follows: atypical glandular cells; atypical glandular cells, favour neoplastic; endocervical adenocarcinoma in situ; and adenocarcinoma.

Advances in the understanding of HPV biology and histological advances were reflected in a revision of TBS in 2014 (Nayar & Wilbur, 2015; Table 4.7). Most of the changes were small, but two major changes were made. In TBS 2014, the cut-off age for reporting benign endometrial cells was changed from 40 years to 45 years. Follow-up studies had reported that the incidence of endometrial carcinoma differed between women in their forties and in their fifties (Weiss et al., 2016; Colletti et al., 2017; Grada et al., 2017; Hinson et al., 2019). In addition, TBS 2014 added chapters covering adjunctive testing, computer-assisted interpretation, education, and risk assessment in cervical cancer (Massad et al., 2013).

Table 4.7 The 2014 Bethesda System for Reporting Cervical Cytology

SPECIMEN TYPE
Indicate conventional smear (Pap smear) vs liquid-based preparation vs other
SPECIMEN ADEQUACY
• Satisfactory for evaluation (describe presence or absence of endocervical/transformation zone component and any other quality indicators, e.g. partially obscuring blood, inflammation, etc.)
• Unsatisfactory for evaluation (specify reason)
o Specimen rejected/not processed (specify reason)
o Specimen processed and examined, but unsatisfactory for evaluation of epithelial abnormality (specify reason)
GENERAL CATEGORIZATION (OPTIONAL)
• Negative for intraepithelial lesion or malignancy
• Other: see Interpretation/Result (e.g. endometrial cells in a woman aged ≥ 45 years)
• Epithelial cell abnormality: see Interpretation/Result (specify squamous or glandular, as appropriate)
INTERPRETATION/RESULT
Negative for intraepithelial lesion or malignancy
When there is no cellular evidence of neoplasia, state this in the General Categorization above and/or in the Interpretation/Result section of the report – whether or not there are organisms or other non-neoplastic findings
Non-neoplastic findings (optional to report)
• Non-neoplastic cellular variations
o Squamous metaplasia
o Keratotic changes
o Tubal metaplasia
o Atrophy
o Pregnancy-associated changes
• Reactive cellular changes associated with:
o Inflammation (includes typical repair)
– Lymphocytic (follicular) cervicitis
o Radiation
o Intrauterine contraceptive device (IUD)
• Glandular cells status post-hysterectomy
Organisms
• *Trichomonas vaginalis*
• Fungal organisms morphologically consistent with *Candida* spp.
• Shift in flora suggestive of bacterial vaginosis
• Bacteria morphologically consistent with *Actinomyces* spp.
• Cellular changes consistent with herpes simplex virus
• Cellular changes consistent with cytomegalovirus
Other
• Endometrial cells (in a woman aged ≥ 45 years)
(Specify if negative for squamous intraepithelial lesion)
Epithelial cell abnormalities
Squamous cell
• Atypical squamous cells
o Of undetermined significance
o Cannot exclude HSIL
• LSIL (encompassing: HPV/mild dysplasia/CIN1)
• HSIL (encompassing: moderate and severe dysplasia, CIS; CIN2 and CIN3)
o With features suspicious for invasion (if invasion is suspected)
• Squamous cell carcinoma

Table 4.7 (continued)

Glandular cell
• Atypical
o Endocervical cells (NOS or specify in comments)
o Endometrial cells (NOS or specify in comments)
o Glandular cells (NOS or specify in comments)
• Atypical
o Endocervical cells, favour neoplastic
o Glandular cells, favour neoplastic
• Endocervical adenocarcinoma in situ
• Adenocarcinoma
o Endocervical
o Endometrial
o Extrauterine
o NOS
Other malignant neoplasms (specify)
ADJUNCTIVE TESTING
Provide a brief description of the test method(s) and report the result so that it is easily understood by the clinician
COMPUTER-ASSISTED INTERPRETATION OF CERVICAL CYTOLOGY
If case examined by an automated device, specify device and result
EDUCATIONAL NOTES AND COMMENTS APPENDED TO CYTOLOGY REPORTS (optional)
Suggestions should be concise and consistent with clinical follow-up guidelines published by professional organizations (references to relevant publications may be included).

CIN, cervical intraepithelial neoplasia; CIS, carcinoma in situ; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion; NOS, not otherwise specified.
Reproduced from Nayar & Wilbur (2015). Copyright 2015, with permission from John Wiley & Sons.

Although the scientific community has made considerable efforts to standardize the criteria for cervical cytology classification, the interpretation of cytology results in substantial variability. For example, in a multicentre RCT designed to evaluate the interpretation of mildly abnormal cytology findings, the reproducibility of monolayer cytological interpretations was moderate (kappa value, 0.46; 95% CI, 0.44–0.48) (Stoler et al., 2001). The disagreement was particularly strong for the ASC-US category, where the concordance was only 42.3%. Studies in Europe also reported disagreement in the ASC-US category (kappa value, 0.10; 95% CI, 0.07–0.13), and the results could not be improved after discussion (kappa value, 0.12; 95% CI, 0.09–0.15) (Ronco et al., 2003). In the first Bethesda Interobserver Reproducibility Study (BIRST-1), 77 images were interpreted by 216 cytotechnologists and 185 pathologists, all of whom were highly experienced, but agreement was obtained for only 67.9% of NILM, 54.1% of LSIL, 22.4% of ASC-H, and 39.9% of ASC-US (Sherman et al., 2007). In the BIRST-2 study for TBS 2014, 518 international participants interpreted 84 digital images (Kurtycz et al., 2017). The overall agreement was 62.8%, which was higher than that in the BIRST-1 study (55.3%). The best agreement was found for NILM (73.4%) and LSIL (86.3%); other results were as follows: 61.7% for ASC-US and 59.5% for HSIL. In a recent study in Brazil, 6536 examinations were reviewed and it was found that kappa values increased from 0.84 to 0.94 (de Morais et al., 2020).

(e) p16/Ki-67 dual staining

The $p16^{INK4a}$ (p16) protein has been widely used in immunocytochemical staining as a biomarker for transforming HPV infection (von Knebel Doeberitz, 2002). The overexpression of p16 in cervical dysplasia is associated with the expression of the E7 oncoprotein of carcinogenic HPV types and can be a surrogate marker of the E7-mediated inactivation of the tumour-suppressor function of the retinoblastoma protein (Schmidt et al., 2011). p16 overexpression is directly connected to cellular transformation by HPV, because E7 expression is required to maintain the phenotype in HPV-associated cancers (von Knebel Doeberitz et al., 1992). p16 overexpression is found in most cervical precancerous lesions and cancers, but it is rarely observed in normal tissue (Klaes et al., 2001).

The expression of the proliferation marker Ki-67 within the same cervical epithelial cell can be used as a surrogate marker of cell cycle deregulation mediated by transforming HPV infection. Although p16/Ki-67 dual staining is independent of morphological interpretation, the interpretation of positive results is operator-independent, not automated. When slides show cervical epithelial cells with brown cytoplasmic p16 immunostaining and red nuclear Ki-67 immunostaining, they could be interpreted as a positive result (Petry et al., 2011). The p16 positivity rate is determined by the distribution of the staining into the cytoplasm or the nucleus and the number of cells that display an overexpression of biomarkers (Tsoumpou et al., 2009). Although the cut-off value varied across the studies, the classification proposed by Klaes et al. (2001) was commonly used. The sensitivity of p16/Ki-67 dual staining using a two-cell cut-off value was nearly equal to that of cytology (82.8% vs 83.8%), but the specificity was higher (62.8% vs 48.7%) (Wentzensen et al., 2005). Although Tsoumpou et al. (2009) reported that the reproducibility of p16 immunostaining is limited because there are insufficient standards for interpretation, recent studies have reported good reproducibility, with kappa values from 0.6 to 0.7 (Stoler et al., 2001; Confortini et al., 2007; Allia et al., 2015; Benevolo et al., 2017). There was no difference in kappa values between experts and non-experts for the interpretation of slides from HPV-positive women (Allia et al., 2015).

p16/Ki-67 dual staining is used for cervical cancer screening, with its use divided into three patterns: primary screening, triage of abnormal cytology, and triage of HPV-positive results. The Primary ASC-US and LSIL Marker (PALM) study was an international collaborative study to evaluate the sensitivity and specificity of p16/Ki-67 dual-stain cytology for primary screening in European countries (Ikenberg et al., 2013). The use of p16/Ki-67 dual staining for primary screening is no longer considered to be an option, because there is a stronger rationale for its use for triage of borderline cytology (ASC-US or LSIL) (Peeters et al., 2019) and, more importantly, of HPV-positive women (Wentzensen et al., 2016; Cuschieri et al., 2018).

In a systematic review, Peeters et al. (2019) compared p16/Ki-67 dual staining with high-risk HPV (hrHPV) testing for triage of ASC-US. The meta-analysis confirmed that p16/Ki-67 dual staining was less sensitive for detection of CIN2+ compared with hrHPV testing (84% vs 93%) but more specific for triage of ASC-US (77% vs 45%). Similar results were obtained when p16 staining was used for triage of ASC-US or when the abnormal cytology threshold was changed to ASC-H (Roelens et al., 2012; Xu et al., 2016).

The sensitivity and specificity of p16/Ki-67 dual staining for women with HPV-positive results were compared with those of cytology, HPV16/18 genotyping, and these methods in combination (Table 4.8). Most studies reported that the sensitivity of p16/Ki-67 dual staining for the detection of CIN2+ was 80–90%. Compared with cytology, the sensitivity of p16/Ki-67 dual staining for the detection of CIN2+ was higher,

Table 4.8 Comparison of performance of p16/Ki-67 dual staining, cytology, and HPV16/18 genotyping for triage of women with HPV-positive results

Reference Country	Outcome: CIN2+						Outcome: CIN3+					
	Sensitivity (%) (95% CI)			Specificity (%) (95% CI)			Sensitivity (%) (95% CI)			Specificity (%) (95% CI)		
	p16/Ki-67 dual staining	Cytology (ASC-US+)	HPV16/18 genotyping	p16/Ki-67 dual staining	Cytology (ASC-US+)	HPV16/18 genotyping	p16/Ki-67 dual staining	Cytology (ASC-US+)	HPV16/18 genotyping	p16/Ki-67 dual staining	Cytology (ASC-US+)	HPV16/18 genotyping
Petry et al. (2011) Germany	91.9 (78.1–98.3)	NA	NA	82.1 (72.9–89.2)	NA	NA	NA	NA	NA	NA	NA	NA
Wentzensen et al. (2012) USA	85.5 (77.8–90.9)	NA	47.6 (38.6–56.7)	59.4 (53.3–65.1)	NA	80.8 (75.5–85.2)	90.6 (73.8–97.5)	NA	75.0 (71.3–80.3)	48.6 (43.5–53.9)	NA	76.1 (71.3–80.3)
Wentzensen et al. (2015) USA	83.4 (77.1–88.6)	76.6 (69.6–82.6)	NA	58.9 (56.2–61.6)	49.6 (46.9–52.3)	NA	86.9 (78.6–92.8)	83.8 (75.1–90.5)	NA	56.9 (54.2–59.5)	48.7 (46.1–51.4)	NA
Gustinucci et al. (2016) Italy	87.6 (75.7~93.6)	77.6 (65.3–86.7)	47.0 (34.0–58.9)	74.9 (69.0–79.0)	72.5 (67.2–77.2)	77.9 (72.8–82.0)	92.3 (74.9–99.1)	96.3 (81.0–99.9)	63.0 (42.4–80.6)	NA	NA	NA
Wright et al. (2017) USA	70.3 (65.3–74.9)	51.8 (46.5–58.3)	NA	75.6 (74.0–77.1)	76.1 (74.6–77.7)	NA	74.9 (69.0–75.7)	51.9 (45.4–58.3)	NA	74.1 (72.5–75.7)	75.0 (73.5–76.5)	NA
Stanczuk et al. (2017) United Kingdom	85.0 (73.4–92.9)	68.3 (55.0–79.7)	61.7 (48.2–73.9)	76.7 (71.1–81.8)	89.1 (84.7–92.7)	70.5 (64.6–76.0)	NA	NA	NA	NA	NA	NA
Wentzensen et al. (2019) USA	88.6 (84.5–92.6)	84.3 (79.7–89.0)	NA	53.1 (51.3–54.9)	42.9 (41.1–44.6)	NA	82.8 (79.4–86.2)	81.1 (77.6–84.7)	NA	55.7 (53.9–57.6)	44.6 (42.9–46.5)	NA
Stoler et al. (2020) USA	NA	NA	NA	NA	NA	NA	86.0	77.2	59.1	60.1	61.6	76.5
Hu et al. (2020) China	63.5 (54.4–71.9)	61.9 (52.8–70.4)	61.9 (52.8–70.4)	85.3 (82.5–87.8)	80.0 (76.9–82.9)	72.4 (68.9–75.6)	64.7 (55.2–73.3)	62.9 (53.5–71.7)	62.9 (53.5–71.7)	84.8 (82.0–87.3)	79.6 (76.5–82.5)	72.1 (68.6–75.3)
Jiang et al. (2020) China	75.0 (50.9–91.3)	NA	NA	50.3 (41.9–58.8)	NA	NA	83.3 (35.9–99.6)	NA	NA	42.7 (34.8–50.8)	NA	NA

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; NA, not available.

Table 4.9 Comparison of guidelines for quality assurance for cytology

Component	Guideline (year published)		
	European Commission (2008)[a]	United Kingdom (2019–2020)[b]	Australia (2018–2019)[c]
Training		✓	✓
Sample collection	✓	✓	✓
Organization (staff, workload)	✓	✓	✓
Material requirement	✓	✓	✓
Quality management	✓	✓	✓
Terminology	✓	✓	✓
Management of abnormal cytology	✓	✓	✓
Follow-up	✓	✓	✓
Laboratory performance indicators	✓	✓	✓
Quality improvement (audit)		✓	✓

[a] Arbyn et al. (2008a).
[b] Public Health England (2019a, b, 2020).
[c] National Pathology Accreditation Advisory Council (2019).
Table compiled by the Working Group.

but the specificity ranged from 50% to 85%; the sensitivity for the detection of CIN3+ was higher, and the specificity was nearly equal. In contrast, the sensitivity of p16/Ki-67 dual staining for the detection of CIN2+ or CIN3+ was always higher than that of HPV16/18 genotyping, but the specificity was lower. Recent studies have reported that combining HPV16/18 genotyping with p16/Ki-67 dual staining increased the sensitivity, with a slight decrease in specificity (Wright et al., 2017; Wentzensen et al., 2019). In Section 4.4.7, the sensitivity and specificity of this combined method for triage of HPV-positive women are compared with those of five other triage methods.

(f) Quality assurance for cytology

Cytological examination depends on the skill and experience of the individual; the interpretation of cervical samples under the microscope is particularly subjective (Arbyn et al., 2008a). The standardization of cytological procedures should always be considered to ensure they are of good quality. Quality assurance should be included in all programmes related to cervical cancer screening, and laboratory management has an important role in quality improvement (Branca & Longatto-Filho, 2015). Continued attention to quality improvement is recommended to ensure that women have access to high-quality screening. Organizational approaches for laboratories include components that address smear-taking, education of both cytotechnologists and cytopathologists, establishment of laboratory quality assurance programmes, management of abnormal cytology, and protocols for follow-up (Farnsworth, 2016). In addition to the European guidelines that established the basic concepts of quality assurance (Arbyn et al., 2008a), guidelines for laboratory quality assurance published in Australia and the United Kingdom also included basic components needed for management and quality improvement (Public Health England, 2019a, b, 2020; National Pathology Accreditation Advisory Council, 2019; Cancer Council Australia Cervical Cancer Screening Guidelines Working Party, 2020) (Table 4.9).

Training for cytotechnologists and cytopathologists is critical for the quality improvement of cytology. For quality assurance of cervical screening, professional accreditation has been

provided on the basis of educational programmes for cytotechnologists. The European guidelines for quality assurance describe the different educational programmes in European countries (Arbyn et al., 2008a). The Australian and United Kingdom guidelines clarify their educational policy and required accreditations for cytotechnologists (National Pathology Accreditation Advisory Council, 2019; Cancer Council Australia Cervical Cancer Screening Guidelines Working Party, 2020; Public Health England, 2020). The National Health Service (NHS) Cervical Screening Programme has also provided educational programmes for smear-takers including general physicians, nurses and midwives (Public Health England, 2016). Continuous training is also required to maintain the quality of interpretation and administration. To harmonize training and develop a quality standard for cervical cancer screening, the Transnational Training Programme in Cervical Cytology (CYTOTRAIN) has produced training materials for cytotechnologists and cytopathologists in Europe (Herbert et al., 2014).

To ensure the accuracy of slide interpretation, cytology laboratories must control the workload of cytotechnologists to help avoid mistakes caused by fatigue or haste (CDC, 1997). Tarkkanen et al. (2003) reported that the annual workload varied among laboratories in Helsinki and the daily average was 30–40 smears. The detection of abnormalities is associated with time spent screening; cytotechnologists with a restricted workload perform better (Renshaw & Elsheikh, 2013). In Australia, the maximum workload for any person involved in primary examination of LBC is 70 slides per day and the hourly workload must not exceed 10 slides (National Pathology Accreditation Advisory Council, 2019). In European countries, the workload limits vary from 25 to 80 slides per hour (Mody et al., 2000). In the USA, federal regulations require workloads to be less than 100 slides per 24 hours (College of American Pathologists, 2014). The American Society of Cytopathology has published quality assurance recommendations for automated screening, including recommendations about productivity and workloads for cytotechnologists (Elsheikh et al., 2013).

Laboratory performance standards for reporting cervical cytology have been established and commonly include rates of unsatisfactory smears, rates of detection of abnormalities, PPVs, and false-negative rates (College of American Pathologists, 2014; National Pathology Accreditation Advisory Council, 2019; Public Health England, 2019a). In the United Kingdom, external quality assessment is defined to assess the performance of cytopathology laboratories and to improve the preparation of LBC slides (Public Health England, 2016).

Quality improvement is an integral component of the management process, and it makes the programmes safe and effective. An audit is the inspection of the quality assurance system to ensure compliance with standards (Branca & Longatto-Filho, 2015). In Australia, a summary of each laboratory's performance standards is submitted annually to the Royal College of Pathologists Quality Assurance Program for collation (Farnsworth, 2016). Laboratories are inspected at least every 3 years and are required to meet these performance measures to claim financial reimbursement. In the United Kingdom, an annual audit programme is carried out to ensure continuous improvement (Public Health England, 2019a).

Some countries have a cytology registry database for quality control and assessment at a national level. In the Netherlands, such a system, the Dutch Network and National Database for Pathology (PALGA), has been in place since 1990 (van Ballegooijen & Hermens, 2000; Casparie et al., 2007). The system has information on all the cytology and pathology results that the laboratories have recorded. In Australia, the state-based Pap test registries collect individual women's cervical cytology and pathology results

from laboratories (Farnsworth, 2016). The system enables direct follow-up in women who receive abnormal results. In Europe, countries collaborate and compare performance indicators of the programmes (Ronco et al., 2009). The low reproducibility of cytology interpretation can be seen when the proportions of abnormal tests, their distribution by grade, and the PPVs are compared among different population-based screening programmes operating in areas with a homogeneous epidemiology of cervical cancer and screening coverage.

Quality assurance systems differ in resource-constrained settings. Cytology screening requires trained personnel and adequate quality control, and quality assurance is frequently insufficient in resource-constrained settings (Gupta et al., 2017). In southern Thailand, Chichareon et al. (2005) found that abnormal cytology detection rates varied from 0.57% to 3.05%. In these areas, pathology laboratories and pathologists were insufficient in number, they underperformed, and the pathologists' roles were not specialized in hospital laboratories. High rates of unsatisfactory samples were reported in conventional cytology (52.3%) and also when LBC was used (47.5%) (Phaliwong et al., 2018). The insufficient follow-up of abnormal smears is also a serious problem. In Thailand, even in the university hospital, 56.1% of women with ASC-US results had colposcopy and 19% could not be followed up (Chichareon & Tocharoenvanich, 2002). Gage et al. (2003) reported a similar experience in Peru, where only 25% of 183 women with an abnormal smear received follow-up. Although cytology is a standardized screening method and the cost is relatively low, the absence of quality control is a major concern.

4.3.2 Beneficial effects of screening using conventional cytology

The 2005 *IARC Handbooks* review evaluated seven cohort studies and 20 case–control studies from multiple countries to review the efficacy of cytology screening in preventing cervical cancer (IARC, 2005). The studies produced consistent evidence of a benefit of cytology-based screening in reducing cervical cancer incidence, which was consistent with the accompanying comprehensive review of ecological trend data in cervical cancer incidence in multiple countries after the introduction of screening. National-level long-term ecological trend data published since the 2005 *IARC Handbook* from multiple countries and world regions continue to support the population-level effectiveness of cytology-based cervical screening (Vaccarella et al., 2013). This is supported by studies in, for example, Brazil (Reis et al., 2020), Canada (Dickinson et al., 2012), Chile (Sepúlveda & Prado, 2005; Pilleron et al., 2020), Europe (Bray et al., 2005; Mendes et al., 2018), the Nordic countries (Vaccarella et al., 2014; Pedersen et al., 2018), the Republic of Korea (Park et al., 2015), Thailand (Sriplung et al., 2014; Virani et al., 2018), Uruguay (Garau et al., 2019), and the USA (Yang et al., 2018). The 2005 *IARC Handbook* noted that the magnitude of the benefit (reduction in disease through screening) was highly variable. The review concluded that the variation in the size of the reduction in risk of cervical cancer through screening was caused largely by variations in the quality of cytology (which will affect its sensitivity) and in programme organization, rather than by measurement error.

The studies published since the 2005 *IARC Handbook* (IARC, 2005) are described and assessed here.

(a) Randomized controlled trials

Only one RCT comparing cytology screening with control conditions (health awareness raising of symptoms and the availability of screening) using incidence and mortality as outcomes has been published (Sankaranarayanan et al., 2005, 2009) (Table 4.10 and Table 4.11). This cluster-randomized trial compared the impact of a single round of screening in four groups (13 clusters per group) – VIA, cytology, HPV testing, and control – in 52 villages in Osmanabad District in Maharashtra state, India. The estimated baseline cervical cancer incidence rate was high, at 20.0 per 100 000 women, with a largely unscreened high-risk population. The study included 131 746 women aged 30–59 years. Of 32 058 women in the cytology group, 25 549 (79.7%) were screened and 1787 (7.0%) had positive results. The PPV for detecting CIN2/3 was 19.3%. Of the 476 women diagnosed with CIN1, 214 were treated (45.0%), and of the 262 women diagnosed with CIN2/3, 234 (89.3%) were treated. During the 8-year follow-up period, cervical cancer developed in 22 of 23 762 women who had negative results on cytological testing (Sankaranarayanan et al., 2009). The diagnosed incidence of cervical cancer in the cytology group was higher than, although not statistically significantly different from, that in the control group (60.7 per 100 000 person-years vs 47.6 per 100 000 person-years; hazard ratio [HR], 1.34; 95% CI, 0.99–1.82). More advanced-stage cancers were diagnosed in the control group than in the cytology group, although this was not statistically significantly different (stage 2 or higher, 23.2 per 100 000 person-years vs 33.1 per 100 000 person-years; HR, 0.75; 95% CI, 0.51–1.10). Mortality from cervical cancer was lower, but not significantly lower, in the cytology group than in the control group (21.5 per 100 000 person-years vs 25.8 per 100 000 person-years; HR, 0.89; 95% CI, 0.62–1.27). [The Working Group noted that the main limitations of the study were that women in the control group were slightly older (mean age, 40 years vs 39 years) (which was adjusted for in the analysis), that screening after health awareness raising in the control group may have minimized the observed impact of screening, and in relation to cytology that a single round was conducted, when it is well established that cytology screening is optimally performed at regular intervals. These results confirmed that even one Pap test can have an impact on incidence of advanced cancers and mortality, but will increase incidence through earlier detection in a medium time period.]

(b) Reviews and meta-analyses

In 2007, the International Collaboration of Epidemiological Studies of Cervical Cancer (ICESCC, 2007) published an analysis of individual-level data collated from 12 observational studies (one cohort study and 11 case–control studies) to analyse risk factors for cervical cancer by type and included history of screening with cytology in the analysis. The analysis included 8097 women with SCC, 1374 women with adenocarcinoma, and 26 445 control women. The women were aged 16–89 years, had not had a hysterectomy, and had had at least one sexual partner. In studies where it was not clear that diagnostic smears had been excluded, only screens 12 months before diagnosis were included. The analysis found that having a past Pap test was associated with a reduced risk of cervical cancer for both SCC (RR, 0.46; 95% CI, 0.42–0.50) and adenocarcinoma (RR, 0.68; 95% CI, 0.56–0.82).

The systematic review and meta-analysis of Peirson et al. (2013) assessed observational cervical screening studies with incidence and mortality as outcomes against unscreened women for the review period of 1995–2012 and published in English or French. The review identified the above-mentioned RCT of Sankaranarayanan et al. (2009) and two cohort studies, one of which was included in the 2005 *IARC Handbooks* review of cytology screening assessing screening

Table 4.10 Basic characteristics of the randomized trial on the efficacy of cervical cancer screening by conventional cytology

Reference	Location	No. of women (screened/ control group)	Accrual period for screening	Age at entry (years)	No. of examinations/ tests in screened/ control group	Incidence of all cervical cancer[a]		Cancer mortality[a]	
						Rate of cervical cancer per 100 000 person-years in screened/ control group	HR (95% CI)[b]	Rate of cervical cancer per 100 000 person-years in screened/ control group	HR (95% CI)[b]
Sankaranarayanan et al. (2005, 2009)	Cluster at village level, India	131 746 eligible women (32 058/31 488)	October 1999 to November 2003	30–59	25 549/1946	60.7/47.6	1.34 (0.99–1.82)	21.5/25.8	0.89 (0.62–1.27)

CI, confidence interval; HR, hazard ratio.

[a] Rates and hazard ratios have been adjusted for age.

[b] Hazard ratios are for the comparison between each intervention group and the control group.

Table 4.11 Results of the randomized trial on the efficacy of cervical cancer screening by conventional cytology

Reference Country	Age at enrolment or screening (years)	Mean duration of follow-up (years)	No. of subjects	Cancer mortality per 100 000 person-years (no. of cancer deaths) in screened/control group	RR (95% CI)
Sankaranarayanan et al. (2009) (see also Sankaranarayanan et al., 2005) India	30–59 Mean age: cytology group, 39; control group, 40	8	Cytology group, 32 058; control group, 31 488	21.5/25.8	0.89 (0.62–1.27)

CI, confidence interval; RR, relative risk.

interval (Herbert et al., 1996). The other cohort study (Rebolj et al., 2009) specifically assessed incidence and mortality in women after negative screens only, not all screened women. The findings of the review are described in the following sections, where included studies are detailed. Only one study overlaps between the meta-analysis of Peirson et al. (2013) and the ICESCC (2007) study: a case–control study of risk factors for cervical cancer from four Latin American countries, reported in two publications (Herrero et al., 1990, 1992).

(c) Cohort studies

Table 4.12 summarizes the results of five cohort studies published since the 2005 *IARC Handbooks* review; two of them focused on older women (Wang et al., 2017; Pankakoski et al., 2019).

Over a period of 12 years, Odongua et al. (2007) followed up 475 398 women aged 30–95 years in the Republic of Korea; the women all had national health insurance and attended a biennial medical examination. Incidence of and death from cervical cancer were assessed using a combination of cancer registry, hospital, and death data records. Estimates were adjusted for age, body mass index, smoking status, alcohol consumption, menarche, and parity. Overall, 57% of women had ever had a Pap test. Compared with screened women with normal screening results, unscreened women had higher incidence of and mortality from cervical cancer (incidence: adjusted RR, 1.12; 95% CI, 1.00–1.25; mortality: adjusted RR, 2.00; 95% CI, 1.37–2.81). Women with abnormal screening results also had higher incidence and mortality than screen-negative women (incidence: adjusted RR, 2.81; 95% CI, 2.54–3.02; mortality: adjusted RR, 2.47; 95% CI, 1.74–3.53). [The Working Group noted that there is insufficient detail in the article to know whether, as seems likely from these findings, diagnostic smears from unscreened women with symptoms were included in the abnormal screening results group. Most studies (see Table 4.12) exclude smears collected in the months before a diagnosis of cervical cancer as evidence of screening and classify women with only these tests as unscreened. If these women are considered as screened, the group of screened women with abnormal results will include unscreened women who develop cancer, biasing the effect of screening overall towards the null. An overall adjusted RR for all screened versus unscreened women is not provided in the study.]

Also in the Republic of Korea, Jun et al. (2009) used data from a national cohort study (the National Health Insurance Corporation Study), which included civil servants and private school employees and their dependents who had health insurance and who participated in at least one routine biennial medical examination between 1995 and 1996. In this study, 253 472 women aged 20 years or older were followed up until 2002 (baseline exclusions were women with previous hysterectomy or cancer; this was not a consent-based study and used routinely collected health information from the insurer). Biennial Pap screening and risk factor surveys were offered by local health services within the cohort, and 52% of women were screened at least once. In total, 241 415 Pap tests were collected, of which 110 were excluded (as diagnostic tests) because they were taken within 3 months of diagnosis of cancer, leaving 241 305 Pap tests. Screening frequency was defined as never, once, or twice or more. Cancer incidence data were taken from the Korean Central Cancer Registry and mortality data for 1995–2002 from the National Statistical Office. After adjustment for age, smoking status, and alcohol consumption, the results showed that women screened twice or more had lower rates of cervical cancer (RR, 0.29; 95% CI, 0.20–0.45), with no significant reduction in those screened only once compared with no screening (RR, 0.90; 95% CI, 0.68–1.18)). Two or more screens were protective against carcinoma in situ of the cervix and across age ranges from

Table 4.12 Cohort follow-up studies on the effectiveness of cervical cancer screening by conventional cytology

Reference Country	Cohort description: no. of women, screening period, source of screening data, and source of follow-up data	Established programme: year of start, screening age, screening interval	Accrual and follow-up periods Person-years	Cervical cancer or precancer end-point, and incidence or mortality age ranges	No. of cases or deaths	Cervical cancer incidence or mortality RR (95% CI)[a]	Adjustments	Comments
Odongua et al. (2007) Republic of Korea	475 398 women with national health insurance aged 30–95 yr who attended a biennial medical examination Screening data from insurance records Incidence and mortality data from national cancer registry, hospital records, and death data records from the National Statistical Office	1992, employees and dependents through national health insurance. In 2000, mandated Pap testing through National Health Insurance Law as part of National Cancer Screening Program	1992–2004 12 yr Person-yr not given	Cervical cancer incidence and mortality Age range at enrolment, 30–95 yr	2523 cases 209 deaths	Incidence: Compared with screened women with normal results (reference) Screened women with abnormal results: 2.81 (2.54–3.02) Unscreened women: 1.12 (1.00–1.25) Mortality: Compared with screened women with normal results (reference) Screened women with abnormal results: 2.47 (1.74–3.53) Unscreened women: 2.00 (1.37–2.81) [Unscreened reference group: Incidence: 0.89 (0.8–1.0) Mortality: 0.5 (0.36–0.73)] [Unscreened reference group: Incidence: 0.36 (0.33–0.39) Mortality: 0.40 (0.28–0.57)]	Age, BMI, smoking status, alcohol consumption, menarche, parity	Only compared unscreened women with screened women with normal results, not all screening. Not clear that diagnostic smears were excluded

Table 4.12 (continued)

Reference Country	Cohort description: no. of women, screening period, source of screening data, and source of follow-up data	Established programme: year of start, screening age, screening interval	Accrual and follow-up periods Person-years	Cervical cancer or precancer end-point, and incidence or mortality age ranges	No. of cases or deaths	Cervical cancer incidence or mortality RR (95% CI)[a]	Adjustments	Comments
Jun et al. (2009) Republic of Korea	253 472 women aged ≥ 20 yr Frequency of Pap testing was determined by the National Health Examination Database Cancer incidence was detected through the Korean Central Cancer Registry and mortality through the National Statistical Office	1988, employees and dependents through national health insurance. In 2000 mandated Pap testing through National Health Insurance Law; biennial cervical cancer screening during the follow-up period	1995–2002 Average follow-up time: 6.5 yr 1 657 130.4 person-yr	Incidence of invasive cervical cancer and CIS of the cervix Age ≥ 20 yr	248 cases of invasive cervical cancer 346 cases of CIS of the cervix	Compared with unscreened women (reference) Incidence of cervical cancer, ≥ 2 screens: 0.29 (0.20–0.45) Incidence of cervical cancer, 1 screen: 0.90 (0.68–1.18) Incidence of CIS of the cervix, ≥ 2 screens: 0.34 (0.25–0.46) Incidence of CIS of the cervix, 1 screen: 0.66 (0.51–0.85)	Age, smoking status, alcohol consumption	Women who were screened ≥ 2 times also had significantly lower rates of all cancers compared with women never screened [supporting the idea of a healthy participant effect]

Table 4.12 (continued)

Reference Country	Cohort description: no. of women, screening period, source of screening data, and source of follow-up data	Established programme: year of start, screening age, screening interval	Accrual and follow-up periods Person-years	Cervical cancer or precancer end-point, and incidence or mortality age ranges	No. of cases or deaths	Cervical cancer incidence or mortality RR (95% CI)[a]	Adjustments	Comments
Dugué et al. (2014) Denmark	1 156 671 women aged 23–51 yr on 1 January 1990 and alive on 31 December 1993 for the 1-round analysis, and 1 030 786 women aged 23–51 yr on 1 January 1990 and alive on 31 December 1997 for the 2-round analysis. Women with gaps in residence in Denmark were excluded In this period, all women were invited to 2 screening rounds, and cytology records were taken from the Danish Pathology Data Bank, National Health Service Register, and National Patient Register The women were followed up until 31 December 2010 (or death or emigration). Follow-up data were from the Danish Civil Registration System and Danish Cause of Death register using Danish unique personal identification numbers	1986, women 23–59 yr, personally invited every 3 yr (90% of women were covered by the guidelines in 1997). Since 2007, every 3 yr for women aged 23–49 yr and every 5 yr for those aged 50–65 yr	1998–2010 Person-yr not given	Mortality due to cervical cancer by screening status as never screened, irregularly screened (attended 1 of 2 rounds), compared with regularly screened (attended both rounds between 1990 and 1997)	No. of cervical cancer deaths: Never screened, 274 Irregularly screened, 152 Regularly screened, 237	Mortality HR compared with regularly screened (1.0) Never screened: 7.91 (6.62–9.46) Irregularly screened: 2.23 (1.81–2.73) [Unscreened reference group: Never screened: 0.13 (0.11–0.15) Irregularly screened: 0.45 (0.37–0.55)]	Adjusted for age by using attained age as time scale in Cox proportional hazards regression	Overall study findings in relation to all-cause mortality: unscreened women had 1.5–2× risk of dying compared with screened women, with a mortality gap maintained over 2 decades. This group also had almost 4× risk of death from other HPV-associated cancers Any cytology test included in screening [this will lead to underestimate of protection from screening]

Table 4.12 (continued)

Reference Country	Cohort description: no. of women, screening period, source of screening data, and source of follow-up data	Established programme: year of start, screening age, screening interval	Accrual and follow-up periods Person-years	Cervical cancer or precancer end-point, and incidence or mortality age ranges	No. of cases or deaths	Cervical cancer incidence or mortality RR (95% CI)[a]	Adjustments	Comments
Wang et al. (2017) Sweden	569 132 women born between 1 January 1919 and 31 December 1945, resident in Sweden since age 51 yr, from the population registry. Women who died or emigrated before age 61 yr or who had invasive cervical cancer or total hysterectomy before age 61 yr were excluded Women entered the cohort at age 61 yr and were followed up until a diagnosis of invasive cervical cancer, a total hysterectomy, emigration from Sweden, age 81 yr, death, or 31 December 2011, whichever came first. Cancer cases were identified through the Swedish National Cancer Registry. National linked registries were used for confounding variables (education level, birth cohort). Screening history was from screening registry	Organized cervical screening programme introduced between 1967 and 1977. Every 3 yr for women aged 23–50 yr and every 5 yr for women aged 51–60 yr. Some areas screening women to age 65 yr	Median follow-up time: unscreened, 10.6 yr; screened, 11.4 yr; overall, 10.9 yr Person-yr not given	Cervical cancer incidence after age 60 yr Data modelled in a competing risk framework (hysterectomy and death as competing events) using screening history at ages 51–60 yr as stratifying variable and first test at age 61–65 yr as exposure of interest Outcome: cervical cancer. Pap tests within 50 d of diagnosis excluded 37% of cohort screened at age 61–65 yr	868 cases of cervical cancer diagnosed at age 61–80 yr	HR for screening at age 61–65 yr stratified by screening status at age 51–60 yr (adjusted for birth cohort, education level) Adequately screened, normal: 0.90 (0.69–1.17) Inadequately screened, normal: 0.82 (0.56–1.22) Unscreened: 0.42 (0.24–0.72) Low-grade abnormality: 0.43 (0.25–0.74) High-grade abnormality: 0.59 (0.36–0.96)	Education level, birth cohort Sensitivity analysis included parity and lifetime diagnosis of COPD as a proxy for smoking status	Extent of benefit from screening women in their 60s varied depending on previous screening history. Provides significant risk reduction for previously unscreened women or women with past abnormalities. Women with normal histories may still benefit from stage shift

Table 4.12 (continued)

Reference Country	Cohort description: no. of women, screening period, source of screening data, and source of follow-up data	Established programme: year of start, screening age, screening interval	Accrual and follow-up periods Person-years	Cervical cancer or precancer end-point, and incidence or mortality age ranges	No. of cases or deaths	Cervical cancer incidence or mortality RR (95% CI)[a]	Adjustments	Comments
Pankakoski et al. (2019) Finland	Cohort of 954 128 women born in 1926–1956 and aged 55–65 yr at the beginning of follow-up, from the population registry Screening history was taken from the screening registry, 1991–2011. Incidence of cervical cancers and deaths in women aged ≥ 55 yr were from the cancer registry Rates were compared with the reference cohort (because uninvited at 65 yr were not Helsinki residents, so had different underlying risk)	Target age 30–60 yr, every 5 yr. Cytology and, since 2012, primary HPV testing has been incorporated into the cervical cancer screening programme. Some use in RCT 2003–2012. Some areas, including Helsinki, invite women to age 65 yr	1991–2014 Median, 11.1 person-yr	Incidence-based mortality risk ratio of cervical cancer for women invited to routine screening at age 65 yr compared with those not invited	No. of cervical cancer deaths: Study cohort (486 869) not invited at age 65 yr, *n* = 212; unadjusted rate, 3.8 per 100 000 Study cohort (59 065) invited (Helsinki) at age 65 yr, *n* = 25; unadjusted rate, also 3.8 per 100 000	Background risk-adjusted RR of death from cervical cancer for women invited at age 65 yr: 0.52 (0.29–0.94), compared with those not invited RR with respect to the uninvited: For women not attending screening: 1.28 (0.65–2.50) For women attending screening: 0.28 (0.13–0.59)	Area of residence (background risk of cervical cancer)	Helsinki area was using cytology. Some areas were using HPV testing Unable to adjust for individual-level hysterectomy

BMI, body mass index; CI, confidence interval; CIS, carcinoma in situ; COPD, chronic obstructive pulmonary disease; d, day or days; HPV, human papillomavirus; HR, hazard ratio; RCT, randomized controlled trial; RR, relative risk; yr, year or years.

[a] Data as reported in source, with conversion to reference group of unscreened women where necessary to standardize comparison.

30 years or older, with no cases recorded in this category in those aged 20–29 years. [The Working Group noted that women who were screened twice or more also had significantly lower rates of all cancers, supporting the idea of a healthy participant effect.]

In Denmark, Dugué et al. (2014) aimed to compare all-cause mortality between cervical screening participants and non-participants and included mortality from cervical cancer as an outcome. Using the Danish registry infrastructure, 1 030 786 women resident in Denmark aged 23–51 years on 1 January 1990 and still alive on 31 December 1997 (a period during which all were offered two rounds of screening) were followed up until death, emigration, or 31 December 2010. The hazard ratio for death from cervical cancer for never-screened women compared with regularly screened women was 7.91 (95% CI, 6.62–9.46) and for irregularly screened women compared with regularly screened women was 2.23 (95% CI, 1.81–2.73).

Two cohort studies focused on older women: in Sweden, Wang et al. (2017) examined the protectiveness of screening against cervical cancer incidence in women older than 60 years, complementing the cohort study of Pankakoski et al. (2019) in Finland, which examined the effectiveness of screening against cervical cancer mortality in women older than 65 years.

Wang et al. (2017) used linked registry databases to follow up 569 132 women in Sweden for a median of 10.9 years and examined their screening history at age 51–60 years to determine the impact of being screened at age 61–65 years on cervical cancer incidence at age 61–80 years. After adjusting for birth cohort and education level, they found that the greatest benefit of screening at age 61–65 years, compared with not screening at that age, was in those women who were unscreened at ages 51–60 years or who previously had abnormalities detected. Hazard ratios were as follows: in unscreened women at age 51–60 years, 0.42 (95% CI, 0.24–0.72); in women with previous low-grade abnormality at age 51–60 years, 0.43 (95% CI, 0.25–0.74); in women with previous high-grade abnormality at age 51–60 years, 0.59 (95% CI, 0.36–0.96). Women with a previous normal history at age 51–60 years had a non-significant reduction in risk through screening at age 61–65 years compared with women with the same history who were not screened. Results were as follows: in women with adequate screening history at age 51–60 years, normal results, 0.90 (95% CI, 0.69–1.17); in women with inadequate screening history at age 51–60 years, normal results, 0.82 (95% CI, 0.56–1.22).

Pankakoski et al. (2019) compared cervical cancer mortality for women in Helsinki offered screening at age 65 years with women from other parts of Finland who were not offered screening at age 65 years but who had been offered routine screening every 5 years from age 30 years to 60 years. The cohort included 954 128 women aged 55–65 years followed up from 1991 to 2011. During the study, most screening was performed using conventional cytology, with small amounts of HPV-based testing during a concurrent RCT. The background risk-adjusted RR of death from cervical cancer for women invited at age 65 years was 0.52 (95% CI, 0.29–0.94), compared with the uninvited. Unsurprisingly, there was an important difference in risk by acceptance of the invitation: for non-attenders, 1.28 (95% CI, 0.65–2.50) and for attenders, 0.28 (95% CI, 0.13–0.59). Self-selection bias may affect these findings (lower-risk women with a history of screening may be more likely to accept the invitation to screen at age 65 years). [The Working Group noted the adequate quality of the study; although women were from different geographical areas, this was adjusted for in the analysis.]

(d) Case–control studies

Peirson et al. (2013) identified 18 case–control studies (one study had four publications) of adequate quality and suitable outcome

measures to estimate the impact of cytology screening on cervical cancer incidence and to consider age range and screening intervals. The data meta-analysis included almost 4800 cases and 18 000 controls from 12 of the studies, and found lower odds of having undergone screening with cytology in women who were diagnosed with cervical cancer (odds ratio [OR], 0.35; 95% CI, 0.30–0.41; $P < 0.00\ 001$) but noted a large degree of heterogeneity. These studies included older data identified through being previously included in two reviews of cervical screening by the United States Preventive Services Task Force. Eleven of these studies were included in the 2005 *IARC Handbooks* review (Aristizabal et al., 1984; Herrero et al., 1992; Sasieni et al., 1996; Hernández-Avila et al., 1998; Jiménez-Pérez & Thomas, 1999; Nieminen et al., 1999; Hoffman et al., 2003; Sasieni et al., 2003) or the 1986 IARC review (Clarke & Anderson, 1979; La Vecchia et al., 1984; Berrino et al., 1986; IARC, 1986). Four additional studies identified by Peirson et al. (2013) (Makino et al., 1995; Talbott et al., 1995; Andrae et al., 2008; Decker et al., 2009), four studies identified but not included in the overall estimate of effect by Peirson et al. (2013) (Zappa et al., 2004; Yang et al., 2008; Sasieni et al., 2009; Kasinpila et al., 2011), and nine studies identified from further literature review (Murillo et al., 2009; Lönnberg et al., 2012; Nascimento et al., 2012; Kamineni et al., 2013; Castañón et al., 2014; Vicus et al., 2015; Rosenblatt et al., 2016; Lei et al., 2019; Wang et al., 2020) are summarized below and in Table 4.13 (web only; available from https://publications.iarc.fr/604); these studies add to the consistency of the literature supporting the effectiveness of cytology-based screening in preventing cervical cancer development and death. Three further case–control studies used mortality as an outcome (Lönnberg et al., 2013; Rustagi et al., 2014; Vicus et al., 2014). The available case–control studies are a mixture of population-based studies using administrative data sets, which avoid participation and recall biases, and studies based on recruitment invitations, which probably suffer from these biases but obtain detailed information to adjust for confounders. Each study has strengths and weaknesses in attempting to estimate the true underlying effect; however, the overall consistency of findings is reassuring, in particular from the studies of Lönnberg et al. (2012, 2013), which examine both incidence and mortality, and attempt to adjust for self-selection bias.

Makino et al. (1995) studied the relationship of screening history with diagnosis of invasive cervical cancer using a case–control design including 198 cases of invasive cervical cancer diagnosed in 1984–1990 in Miyagi, Japan, each matched with two controls by age and area. They divided the cases into those that were detected by screening, who were assigned controls from screening programme records, and those that were diagnosed as outpatients, who were matched with other gynaecological outpatients. They determined ever-screened status using programme records or, if a woman reported on a questionnaire that she was screened elsewhere, accepted self-report. They excluded women with a history of abnormal screening results; it is unclear whether this exclusion applies to both cases and controls and the impact it will have on the correct assignment of whether a woman has ever been screened compared with the underlying population. They found a protective OR of 0.14 (95% CI, 0.088–0.230) for ever being screened, consistent across the age ranges 34–49 years and 50–74 years. [The Working Group noted that the limitations of this study – the exclusion of women with abnormal screening results and the acceptance of self-report – may have resulted in an overestimate of the true effect of screening.]

Talbott et al. (1995) examined self-reported screening history from cases of invasive cervical cancer sourced from the Pennsylvania Cancer Registry and age- or area-matched controls. Because screening history was obtained from consent-based interviews up to 2 years after

diagnosis, only 143 women (30% of cases) with a matched control were included in the final analysis (ages 25–79 years), resulting in cases with an earlier stage of disease than the source sample. Although it acknowledged both selection bias and likely recall bias, the study estimated an OR of no Pap test in the previous 3 years of 3.10 (95% CI, 1.45–6.64), adjusted for smoking status, marital status, income, physician's visit within 3 years, number of pregnancies, age at first pregnancy, number of long-term relationships, use of birth control, and use of condoms. [The Working Group noted that the findings should be interpreted with caution because of the poor participation rate of cases; cases with advanced disease at diagnosis were systematically underrepresented.]

Zappa et al. (2004) examined the screening history of 208 cases of invasive cervical cancer in women aged 70 years or younger at diagnosis between 1994 and 1999 and 832 age-matched controls in Florence, Italy. The study aimed to assess the impact of screening on the incidence of adenocarcinoma compared with squamous cancers, and the impact of screening by age in women younger than or older than 40 years. High-grade CIN and cancers were identified through the Tuscany Tumour Registry, and screening history was collected from a computerized archive estimated to contain about two thirds of all the screening tests in the area. Smears taken in the 12 months before the index date of the case were excluded. Four randomly selected controls with no record of hysterectomy and who were resident for at least 5 years in the area per case (matched on year of birth) were selected from the municipality residence database. After adjustment for civil status and birthplace, screening was found to be protective against cervical cancer (< 3 years since last test: OR, 0.25; 95% CI, 0.15–0.42; 3–< 6 years since last test: OR, 0.34; 95% CI, 0.21–0.56); ≥ 6 years since last test: OR, 0.56; 95% CI, 0.38–0.82). However, no significant protection was observed for adenocarcinomas alone (< 3 years since last test: OR, 0.65; 95% CI, 0.26–1.65), and women older than 40 years had stronger and more consistent protection against SCCs over time from screening.

Andrae et al. (2008) assessed all 1230 invasive cervical cancer cases diagnosed in Sweden between 1999 and 2001 against the screening history in the previous 6 years of five population-based age-matched controls per case (6124 total). All data were obtained from population-based linked data registries, avoiding recall or selection bias. Women who had not been screened in the recommended interval for their age had higher odds of cervical cancer (OR, 2.52; 95% CI, 2.19–2.91), with consistent findings across age groups. Screening was also protective against non-SCC cancers (SCC: OR, 2.97; 95% CI, 2.51–3.50; non-SCC: OR, 1.59; 95% CI, 1.20–2.11).

Yang et al. (2008) undertook a case–control study in New South Wales, Australia, where biennial cytology screening was recommended for women aged 20–69 years. Data on 877 cases diagnosed with invasive cervical cancer between 2000 and 2003 were obtained from the cancer registry and controls from the Pap Test Register, which contains almost all screening results. However, to have a record in the Pap Test Register a woman needs to have been screened at least once. [The Working Group noted that this may have led to the 2614 age-matched controls being more likely to have been screened than the general population from which the cases were drawn, which could bias estimates in favour of screening being protective. Therefore, the study findings are applicable to screened women rather than to the general population.] The exposure of interest was screening in the 4-year period before diagnosis, and results were adjusted for the result of the first Pap test in the previous 6 years. Compared with no screening in the previous 4 years, irregular screening had an OR of 0.189 (95% CI, 0.134–0.265) and regular screening

had an OR of 0.065 (95% CI, 0.044–0.096). If restricted only to cases with any screening history on the screening registry, to match selection criteria with controls, estimates were attenuated somewhat: irregular screening OR, 0.215 (95% CI, 0.150–0.309), regular screening OR, 0.070 (95% CI, 0.046–0.106). Results were consistent across 10-year age groups and for both SCC and non-SCC cancers.

In Manitoba, Canada, Decker et al. (2009) compared screening in the previous 5 years from administrative claims between 666 cervical cancer cases aged 18 years or older notified to the cancer registry in 1989–2001 and 3343 age- and area-matched controls (5 per case) sourced from a state-wide universal health insurance register. Women who had not had a Pap test in the previous 5 years had higher odds of cervical cancer (OR, 2.77; 95% CI, 2.30–3.30).

In a case–control study in four areas of Colombia, Murillo et al. (2009) enrolled 200 cases aged 25–69 years from pathology records and 200 age- and neighbourhood-matched controls. Screening history was compiled using blinded review, excluding diagnostic smears, and nurses conducted structured risk factor interviews. After adjustment for age at first intercourse, age at first birth, parity, use of oral contraceptives, number of sexual partners, insurance status, and literacy, the OR for cervical cancer in women who had no screening in the previous 36 months was 3.54 (95% CI, 2.01–6.24).

Sasieni et al. (2009) described findings by histological type using their previous population-based case–control study in a data audit of women aged 20–69 years using the routine cytology database in the United Kingdom (Sasieni et al., 1996, 2003). Using data from 3305 cases and 6516 controls, they found that screening within 10 years of diagnosis provided greater protection against SCC and adenosquamous cancers than against adenocarcinoma (adenocarcinoma: OR, 0.72; 95% CI, 0.54–0.95; SCC: OR, 0.37; 95% CI, 0.32–0.41; adenosquamous cancer: OR, 0.25; 95% CI, 0.15–0.43).

In a hospital-based case–control study in Thailand, Kasinpila et al. (2011) compared 130 women aged 30–64 years diagnosed with invasive cervical cancer in four tertiary hospitals with age-matched controls who were patients or visitors at the same hospitals. Screening history and risk factor information were collected by structured interview. After adjusting for age at first intercourse, alcohol consumption, and use of oral contraceptives, they found that any number of tests more than 6 months before the diagnosis date was protective (for 1–5 tests: OR, 0.45; 95% CI, 0.25–0.84; for ≥ 6 tests: OR, 0.29; 95% CI, 0.11–0.82) and that more recent tests were more protective (test in previous 1–2 years: OR, 0.27; 95% CI, 0.13–0.56; test ≥ 3 years ago: OR, 0.42; 95% CI, 0.20–0.88).

The study of Lönnberg et al. (2012) in Finland compared screening in 1546 cervical cancer cases and 9276 age-matched controls using cancer registry, screening registry, and population registry data to avoid selection and recall biases. A statistical adjustment was made to correct for self-selection bias. The estimated association between cervical cancer and screening participation was significant across stage and cancer types (OR, 0.53; 95% CI, 0.46–0.62) and was statistically significant in the individual 5-year age bands between the ages of 40 years and 64 years and in the 15-year age bands of 40–54 years (OR, 0.44; 95% CI, 0.35–0.56) and 55–69 years (OR, 0.37; 95% CI, 0.27–0.52), with a smaller impact in the 25–39 year age group (OR, 0.81; 95% CI, 0.63–1.05).

In a hospital-based case–control study in Rio de Janeiro, Brazil, Nascimento et al. (2012) compared 152 cases with 169 age- and area-matched controls who were visitors to the same hospital. The researchers used a consent-based model and comprehensive risk factor survey to gather screening and other history, recruiting 152 of 169 (89.8%) of eligible cases aged 25–68 years,

90% of whom had SCC. After adjustment for education level, age, municipality, and tobacco use, it was found that reporting three or more Pap tests 3 years before the index date was associated with a lower odds of cervical cancer (OR, 0.16; 95% CI, 0.074–0.384).

Kamineni et al. (2013) assessed the effectiveness of screening women aged 55–79 years in a case–control study in the USA involving 69 cases of invasive cervical cancer and 208 age-matched controls. Women were members of one of two large health insurers, and screening and medical or demographic history for 7 years before the case diagnosis date was obtained through medical record review. After adjustment for age and smoking status, the OR for cervical cancer in those screened 1 year previously (estimated duration of occult phase) was 0.23 (95% CI, 0.11–0.44). The greatest reduction in risk was observed in the year after screening; the incidence returned to that in unscreened women 5–7 years after a negative screen test result.

In the accompanying case–control study of the impact of screening on cervical cancer mortality, Lönnberg et al. (2013) analysed the screening history of 506 women who died in the period 2000–2009 and 3036 age-matched population-based controls. After adjustment for self-selection bias, the results showed a protective effect of an index screen (defined as the last age group invitation and possible screening test within the 66 months before the diagnosis), with an OR of 0.34 (95% CI, 0.14–0.49). No protective effect on mortality from adenocarcinoma was detected, and the effect on mortality was lowest for those aged 25–39 years (OR, 0.70; 95% CI, 0.33–1.48).

Castañón et al. (2014) conducted a population-based case–control study in England and Wales to consider the effect of screening women aged 50–64 years on the incidence of cervical cancer in women aged 65 years or older. The study included 1341 cases diagnosed between 2007 and 2012 and 2646 age-matched controls (two per case, including one from the same general practice). Screening with an interval of < 5.5 years compared with no screening in women aged 50–64 years resulted in an OR for cervical cancer after age 65 years of 0.25 (95% CI, 0.21–0.30). Protection decreased with time since last screen, and the estimated absolute risks over time for the population who were screened at age 50–64 years supported the conclusion that there was low risk in women with adequate negative screening and justified cessation of screening at age 65 years for this group.

Rustagi et al. (2014) conducted a case–control study in the USA in health-care enrolees aged 55–79 years, to assess the effect of screening on cervical cancer mortality in older women. Women who had died from cervical cancer between 1980 and 2010 ($n = 39$) were matched to two controls each ($n = 80$) by health plan, age, and duration of health plan enrolment. Screening in the 7 years before the index date was protective against cervical cancer death (OR, 0.26; 95% CI, 0.10–0.63) after adjustment for matching characteristics, smoking status, marital status, and race or ethnicity.

Vicus et al. (2014) analysed the mortality from cervical cancer and the effectiveness of cytology screening by age group in 1052 cases and 10 494 controls aged 20–69 years diagnosed between 1998 and 2008 in Ontario, Canada. State-wide administrative data sets were used to obtain screening history and to obtain age-matched, income-matched controls, and cases were identified from the cancer registry. Screening 3–36 months before the date of diagnosis was found to be protective in all age groups 30 years or older (ORs from 0.28 to 0.60). In a related analysis of incidence, using 5047 cases and 10 094 controls, Vicus et al. (2015) detected a significant protective effect of screening 3–36 months before the date of diagnosis only in the age groups 40–44 years (OR, 0.82; 95% CI, 0.69–0.97), 50–54 years (OR, 0.59; 95% CI, 0.48–0.73), 55–59 years (OR, 0.52; 95% CI,

0.48–0.73), and 60–64 years (OR, 0.59; 95% CI, 0.46–0.76).

Rosenblatt et al. (2016) examined the effect of cervical screening from age 65 years for up to 7 years between 1991 and 1999 in a population from 11 areas of the USA, using Medicare insurance claims data and Surveillance, Epidemiology, and End Results cancer registry data. The study identified 1267 cases, and these were matched to 10 137 controls (up to 8 controls per case) on age and geographical location. Data on previous hysterectomy were not available for controls, but population-based data were used to estimate the effect on risk of removal of hysterectomized controls. After adjustment for race and postal code-level income, the results suggested that having a Pap test 2–7 years before diagnosis provided significant protection against cervical cancer (OR, 0.64; 95% CI, 0.53–0.78). After adjustment also for the likely prevalence of hysterectomy in controls, the protective effect of screening increased (OR, 0.38; 95% CI, 0.32–0.46). Effectiveness was seen across the age range but was greatest in women aged 65–74 years (hysterectomy-adjusted OR, 0.24; 95% CI, 0.15–0.37), women aged 75–84 years (hysterectomy-adjusted OR, 0.44; 95% CI, 0.34–0.55), and women aged 85–100 years (hysterectomy-adjusted OR, 0.44; 95% CI, 0.29–0.66). In women aged 72 years and older who had complete exposure data for the ascertainment period 1991–1999, the greatest effects were seen in preventing squamous carcinoma (hysterectomy-adjusted OR, 0.31; 95% CI, 0.23–0.40), regional disease (hysterectomy-adjusted OR, 0.27; 95% CI, 0.20–0.39), and distant disease (hysterectomy-adjusted OR, 0.30; 95% CI, 0.16–0.58). [The Working Group noted that the main limitation of this study is that the determinants of screening participation in this age group in this setting are not known. Routine screening was not recommended in previously screened older women during this period, although 3-yearly screening was funded by Medicare. Previous screening history before age 65 years was not available. The results may therefore not be applicable to a general population for which routine screening is recommended.]

Lei et al. (2019) conducted a population-based nested case–control study in Sweden using the linked population registry infrastructure to examine whether cytology screening has a protective effect on the incidence of adenosquamous cancer and rare types of invasive cervical cancer (RICC) (e.g. clear cell carcinoma, large cell carcinoma, glassy cell carcinoma, neuroendocrine carcinoma). Cases of invasive cervical cancer diagnosed in Sweden in 2002–2011 were identified from the Swedish Cancer Registry and underwent clinical and histopathological review, which resulted in the identification of 338 cases of adenosquamous cancer (49%) and RICC (51%). For each case, 30 controls without hysterectomy or history of cervical cancer and who were alive and living in Sweden at the date of diagnosis of the case were selected from the total population register using incidence density sampling and matched on year of birth. Cervical screening data from the previous two screening rounds (women aged 30 years or older were included to enable two screening rounds) were obtained from the national screening registry, and tests within 6 months of the date of diagnosis of the case were excluded. ORs were interpreted as incidence rate ratios. After adjustment for education level, two screening tests compared with none was associated with a substantially lower risk of adenosquamous cancer (IRR, 0.22; 95% CI, 0.14–0.34) and RICC (IRR, 0.34; 95% CI, 0.21–0.55). Protection was greatest for those aged 30–60 years, for adenosquamous cancers, with two tests compared with one, and against more advanced cancers. Protection was seen for both HPV-positive and HPV-negative cancers and across rare cancer types.

Wang et al. (2020) undertook an audit of the Swedish cervical screening programme and presented a population-based nested case–control analysis of cervical cancer risk by screening

status. The authors used the same methods as Lei et al. (2019) but included all cervical cancer cases (n = 4254) and 120 006 controls. Women aged 26–28 years had one screening round examined. Women with no screening tests compared with women who had been screened in the last two rounds had an OR of 4.1 (95% CI, 3.8–4.5) for cervical cancer. Attending one of the two last screens only lowered the odds ratio somewhat (women who missed the last screening round but attended the screening round before: OR, 2.4; 95% CI, 2.2–2.7; women who attended the last screening round but missed the one before: OR, 1.6; 95% CI, 1.5–1.8).

(e) Screening intervals and age range for screening

The Peirson et al. (2013) meta-analysis examined the evidence from 14 studies, including two cohort studies (Herbert et al., 1996; Rebolj et al., 2009) and 12 case–control studies (La Vecchia et al., 1984; Berrino et al., 1986; Herrero et al., 1992; Makino et al., 1995; Sasieni et al., 1996, 2003, 2009; Jiménez-Pérez & Thomas, 1999; Hoffman et al., 2003; Miller et al., 2003; Zappa et al., 2004; Andrae et al., 2008; Yang et al., 2008; Kasinpila et al., 2011), to review screening intervals for protection against incident cervical cancer. The meta-analysis also included four studies that considered ages of commencement and cessation of screening: three case–control studies (Sasieni et al., 1996, 2003, 2009; Hoffman et al., 2003; Andrae et al., 2008) and one cohort study (Rebolj et al., 2009). Differences in study designs prevented any pooling of data to analyse screening intervals, but the review had four key consistent findings: (i) the shortest time interval since the last screen in each study consistently had the highest degree of protection associated with it, (ii) screening intervals of 5 years or less consistently appear to offer protection, (iii) longer intervals between screens provide diminishing protection, but (iv) any history of screening is more protective than no history of screening. No data pooling was possible in examining ages of commencement and cessation of screening. The evidence suggested that screening in women younger than 30 years may be less effective, but evidence is strong for a beneficial effect in women older than 30 years, including in women aged 65 years or older. The more recent data reviewed above support these conclusions that more recent screening confers greater protection, that screening in women younger than 30 years may be of more limited benefit (Lönnberg et al., 2012, 2013; Vicus et al., 2014, 2015), and that there is evidence for the effectiveness of screening older women, noting that women who have not been screened regularly, or who have had previous abnormal screening results, are likely to benefit most from screening at older ages (Kamineni et al., 2013; Castañón et al., 2014; Rustagi et al., 2014; Rosenblatt et al., 2016; Wang et al., 2017; Pankakoski et al., 2019).

4.3.3 Beneficial effects of screening using LBC

(a) Accuracy of LBC compared with conventional cytology

Several systematic reviews and meta-analyses have been published providing estimates of the sensitivity and specificity of LBC and comparing the sensitivity, specificity, and PPV of LBC systems with those of conventional cervical testing in terms of their ability to identify biopsy-confirmed CIN2 or CIN3 (Austin & Ramzy, 1998; Payne et al., 2000; Bernstein et al., 2001; Sulik et al., 2001; Davey et al., 2006; Arbyn et al., 2008b; Whitlock et al., 2011; Chen et al., 2012; Fokom-Domgue et al., 2015; Mustafa et al., 2016). Both techniques are based on the same principles to identify precancerous lesions, using the same staining and interpretation methods and almost identical sampling methods.

Most early studies used a paired-sample design, with either split samples or direct-to-vial sampling. In the split-sample method, the conventional slide is made first, and then the

brush and/or spatula is rinsed in the medium for LBC to collect the remaining cells. In the direct-to-vial sampling method, a dedicated sample is collected for LBC by rinsing the spatula and/or brush in the vial containing the liquid medium; a separate sample for conventional cytology is taken before or after the LBC sample. Both methods may introduce some biases. For example, in split samples, the LBC component, which uses the residual sample after smearing for the conventional slide, systematically starts with less cellular material. In direct-to-vial studies, samples for conventional cytology and LBC are taken separately, and if the two samples are taken close together in time, the second sample will take cells from a cervix that has already been scraped, possibly with less cellular material and a higher probability of bleeding, whereas if the two samples are taken at distant time points, they could reflect different conditions of the cervix (i.e. the lesions could evolve or new lesions could emerge) (Cheung et al., 2003; Colgan et al., 2004; Fremont-Smith et al., 2004). Randomizing the order of sampling could avoid this bias.

Most early studies included relatively small numbers of women, and in order to have enough statistical power to estimate sensitivity, they could not recruit samples from the screening population but needed to include in their study population more women with CIN2+, usually including those referred for colposcopy. This selection may introduce a bias by selecting women who had a recent positive test with the technique used at that time in the screening programme (usually conventional cytology), thus overestimating both conventional cytology true-positive and false-positive results, as was discussed by some authors of these early studies (Confortini et al., 2004). Under certain conditions, these studies could accurately estimate sensitivity and, with the limitation explained below, specificity, but they could not estimate the referral rate that would be experienced in a screening population and consequently the PPV. When using a cytology positivity threshold of ASC-US or worse or LSIL or worse, the cytologist is looking for the cytological signs of a risk factor for the clinically relevant lesions (i.e. HPV infection) and not only for the lesion itself (i.e. CIN2+). Consequently, the test is also dependent upon the underlying prevalence of HPV infection in the tested population for its accuracy. In particular, the specificity of the test decreases when the prevalence of HPV infection increases (Giorgi Rossi et al., 2012).

The quality of the primary studies varied, and most studies had methodological deficiencies and inadequate follow-up (Nanda et al., 2000; Sulik et al., 2001; Davey et al., 2006). In particular, in their systematic review Davey et al. (2006) found that studies of high methodological quality with lower risk of bias estimated very similar sensitivities for LBC and conventional cytology, whereas low-quality studies estimated slightly higher sensitivity for LBC. Similarly, Arbyn et al. (2008b) estimated a pooled sensitivity for LBC of 90.4% (95% CI, 82.5–95.0%) when ASC-US was the threshold and 79.1% (95% CI, 70.1–86.0%) when LSIL was the threshold. For conventional cytology, the pooled sensitivity was 88.2% (95% CI, 80.2–93.2%) when ASC-US was the threshold and 75.6% (95% CI, 66.5–83.0%) when LSIL was the threshold. Therefore, the relative sensitivity estimate for LBC versus conventional cytology was close to 1: 1.03 (95% CI, 0.97–1.09) for an ASC-US threshold and 1.03 (95% CI, 0.96–1.11) for an LSIL threshold. Specificity was higher for conventional cytology when ASC-US was used as the threshold (relative specificity LBC vs conventional cytology, 0.91; 95% CI, 0.84–0.98) and similar when LSIL was used as the threshold (relative specificity LBC vs conventional cytology, 0.97; 95% CI, 0.94–1.01). In their systematic review on HPV test accuracy, Koliopoulos et al. (2017) produced estimates of the absolute sensitivity and specificity of conventional cytology and LBC in studies where cytology was compared with HPV testing. In this review, both cytological methods had lower

sensitivity compared with previous studies: when ASC-US was used as the test threshold, the pooled sensitivity for conventional cytology was 65.9% (95% CI, 54.9–75.3%) for the detection of CIN2+ and 70.3% (95% CI, 57.9–80.3%) for the detection of CIN3+; with the same threshold, the pooled sensitivity for LBC was 75.5% (95% CI, 66.6–82.7%) for the detection of CIN2+ and 70.3% (95% CI, 57.9–80.9%) for the detection of CIN3+. However, the pooled specificity was higher for conventional cytology than for LBC. [To estimate the absolute sensitivity and specificity, colposcopic assessment is required for all subjects to confirm histological diagnosis as a reference standard (Branca & Longatto-Filho, 2015), and because this recent systematic review included studies without systematic assessment of all women, verification bias could not be completely excluded (Fokom-Domgue et al., 2015; Mustafa et al., 2016; Koliopoulos et al., 2017). Furthermore, these estimates come from different studies for conventional cytology and LBC, so the estimates cannot be directly compared.]

Larger studies in low-risk populations, often nested in routine screening programmes, started in the first decade of the 2000s. Some of these studies used a paired-sample design, mostly split samples (Coste et al., 2003; Almonte et al., 2007; Davey et al., 2007; Halford et al., 2010; Tanabodee et al., 2015); others were controlled trials, either individually randomized (Obwegeser & Brack, 2001; Ronco et al., 2007; Maccallini et al., 2008; Sykes et al., 2008) or cluster-randomized (Taylor et al., 2006; Strander et al., 2007; Siebers et al., 2009; Klug et al., 2013). Finally, others were pilot population-based studies with historical or concurrent non-randomized controls (Beerman et al., 2009; Akamatsu et al., 2012; Sigurdsson, 2013; Rebolj et al., 2015; Rozemeijer et al., 2016, 2017; Ito et al., 2020).

(b) Evidence on relative detection and relative PPV from RCTs

In an RCT, the target population is divided into two groups, whose background is expected to have the same characteristics, aside from random fluctuations (Ronco et al., 2007). In large population-based randomized studies, usually only women with a positive test result are assessed. It is therefore impossible to compute absolute sensitivity and specificity. Nevertheless, in this setting, relative detection is a correct estimator of relative sensitivity, and relative referral rate for assessment and relative PPV measure how the specificity of the two tests affects screening efficiency.

Eight RCTs were conducted (Table 4.14) with varying test thresholds and outcomes; seven reported results using ASC-US as the test threshold (Obwegeser & Brack, 2001; Taylor et al., 2006; Ronco et al., 2007; Strander et al., 2007; Maccallini et al., 2008; Sykes et al., 2008; Siebers et al., 2009), and four reported data for an LSIL threshold (Taylor et al., 2006; Ronco et al., 2007; Strander et al., 2007; Klug et al., 2013; Table 4.15).

In a study in a high-risk population in South Africa, Taylor et al. (2006) included colposcopic assessment for all women, which enabled the estimation of the absolute sensitivity and specificity for conventional cytology and LBC. The authors calculated the sensitivity and specificity for conventional cytology and LBC. The sensitivity of conventional cytology for the detection of CIN2+ was 83.6% (95% CI, 71.2–92.2%), with a specificity of 85.1% (95% CI, 83.6–86.5%); the sensitivity of LBC for the detection of CIN2+ was 70.6% (95% CI, 58.3–81.0%), with a specificity of 84.8% (95% CI, 83.5–86.1%).

The only other RCT with colposcopic assessment for all women was conducted in New Zealand (Sykes et al., 2008). In this study, women referred to a colposcopy clinic were randomized to LBC or conventional cytology. The study cannot give information on referral and PPV,

but gave a rather precise estimate of the relative sensitivity: 1.0 (95% CI, 0.83–1.21). [The Working Group noted a low risk of bias in this study.]

The study by Obwegeser & Brack (2001) in Switzerland recruited women of any age attending gynaecology services for opportunistic screening, including women in age ranges for which screening is not recommended. These findings should be interpreted with caution because the only published report included only the assessment of women with high-grade cytological lesions, whereas assessment of women with ASC-US and LSIL was not yet available. LBC classified a higher proportion of women as having LSIL (4.7%) than did conventional cytology (3.7%). The authors found no effect on sensitivity. [The Working Group noted a high risk of bias in this study.]

The study of Ronco et al. (2007) in Italy randomized women to LBC plus HPV testing or to conventional cytology. The study also enabled a comparison between the baseline results for LBC alone versus conventional cytology, because the LBC reading was performed blinded to the HPV test result, although colposcopy was not performed blinded to the HPV test result, which could be expected to increase the index of suspicion for the colposcopist. When the ASC-US threshold was used, the study found a small, non-significant increase in the CIN2+ detection rate using LBC, but not in the CIN3+ detection rate, and the PPV was much lower with LBC than with conventional cytology. When the LSIL threshold was used, LBC had a non-significantly lower detection rate and a similar PPV. [The Working Group noted some concern of bias in this study.]

The study of Strander et al. (2007) in Sweden allocated women to LBC or conventional cytology by randomization of the week of the scheduled appointment. The outcome (CIN2+) was assessed with passive follow-up through the pathology registry, without knowing how the women were individually managed. A 60% increase in detection using LBC was found, with a similar PPV. However, the results should be interpreted with caution given that some imbalance in randomization occurred, because adjusting for age and screening centre produced substantially different ORs compared with unadjusted figures. [The Working Group noted some concern of bias in this study.]

A small RCT in Italy (Maccallini et al., 2008) found no difference in either relative detection or relative PPV but reported strong heterogeneity between centres for relative PPV. The authors noted a higher compliance to colposcopy in the LBC group than in the conventional cytology group, and adjustment for non-compliance reduced the difference in detection between the two groups. [The Working Group noted some concern of bias in this study.]

The largest RCT was conducted in the Netherlands and randomized about 90 000 women (Siebers et al., 2009). The study raised no concerns about randomization and ascertainment procedures, and the sample size enabled precise estimates to be obtained. The authors found similar detection rates for CIN2+ and CIN3+ (CIN2+ relative detection rate, 1.00; 95% CI, 0.84–1.20; CIN3+ relative detection rate, 1.05; 95% CI, 0.86–1.29) and similar PPVs (relative PPV, 0.99; 95% CI, 0.80–1.22) in the two groups. [The Working Group noted a low risk of bias in this study.]

Klug et al. (2013) randomized 20 practices in Germany to use LBC or conventional cytology. The study also included the use of computer-assisted technology in addition to LBC, but results were given separately for manual reading and computer-assisted reading. Nevertheless, the use of computer-assisted reading was used to centralize LBC reading in one laboratory, and conventional cytology was read in nine different laboratories. In Germany the standard cytology classification is the Munich II nomenclature (Hilgarth, 2001). This is the only RCT that reported a more than 2-fold increase in detection rate with LBC

Table 4.14 Study characteristics of randomized controlled trials comparing cervical cancer screening by liquid-based cytology versus conventional cytology

Reference Trial, country	Randomization	No. of women	Population	Age at entry (years)	LBC procedure	Reference standard	Blinding of histological assessment?	Reported end-points	Long-term outcomes
Obwegeser & Brack (2001) Switzerland	Individual	Conv.: 1002 LBC: 997	Opportunistic screening	15–≥ 70	ThinPrep 2000	Colposcopy for women with HSIL cytology; for ASC-US and LSIL, follow-up was mostly incomplete	No	CIN2+	NR
Taylor et al. (2006) South Africa	No; practice rotating every 6 mo	Conv.: 2444 LBC: 3114	High-risk population	35–65	ThinPrep 2000	Colposcopy for all women	Yes	CIN2+ CIN3+	Not possible. Women were all referred for colposcopy
Ronco et al. (2007) NTCC, Italy	Individual	Conv.: 22 466 LBC: 22 708	Screening	25–60	ThinPrep	Colposcopy for all positive	CIN reviewed blindly	CIN2+ CIN3+	Not possible. Women were managed according to HPV test results
Strander et al. (2007) Sweden	Randomized per week of appointment	Conv.: 8810 LBC: 4674	Screening	23–60	ThinPrep 2000	Referral as routine practice; histology searched through registries	Yes	CIN2+	Cumulative incidence up to 3 yr and 7 mo
Sykes et al. (2008) New Zealand	Individual	Conv.: 453 LBC: 451	Women in colposcopy clinics	16–75	SurePath	Colposcopy-guided biopsy	No	CIN2+	NR
Maccallini et al. (2008) Italy	Individual	Conv.: 4299 LBC: 4355	Screening	25–64	ThinPrep	Colposcopy for all positive	No	CIN2+ CIN3+	NR
Siebers et al. (2008, 2009) NETHCON, Netherlands	Cluster RCT; family practice as randomization unit	Conv.: 40 562 LBC: 49 222	Screening	25–60	ThinPrep 3000	Referral as routine practice. All follow-up tests blindly reviewed	Yes	CIN2+ CIN3+	NR
Klug et al. (2013) Germany	Randomized per week of visit	Conv.: 9352 LBC: 11 555	Opportunistic screening	≥ 20	ThinPrep with/without Imaging System	Colposcopy for all women with LSIL+	No	CIN2+ CIN3+	NR

ASC-US, atypical squamous cells of undetermined significance; CIN, cervical intraepithelial neoplasia; Conv., conventional cytology; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion; LBC, liquid-based cytology; LSIL, low-grade squamous intraepithelial lesion; mo, month or months; NETHCON, Netherlands ThinPrep versus Conventional Cytology Trial; NTCC, New Technologies for Cervical Cancer Screening; NR, not reported; RCT, randomized controlled trial; yr, year or years.

Table 4.15 Comparison of test performance between liquid-based cytology and conventional cytology in randomized controlled trials

Reference Country	Age (years)	Threshold	Total number		Detection rate (%)		PPV (%)	Unsatisfactory cytology
					Outcome: CIN2+	Outcome: CIN3+	Outcome: CIN2+	
			Conv.	LBC	RR (95% CI)[a]	RR (95% CI)	RR (95% CI)[a]	RR (95% CI)
Obwegeser & Brack (2001) Switzerland	15–≥ 70	ASC-US	1002	997	0.92 (0.41–2.07)[b]	NA	NA	NA
Taylor et al. (2006) South Africa	35–65	ASC-US	2444	3114	0.81 (0.54–1.21)	0.67 (0.39–1.14)	0.83 (0.56–1.21)	2.85 (1.72–4.72)
Strander et al. (2007) Sweden	23–60	LSIL	8810	4674	1.63 (1.09–2.43)	NA	1.02 (0.75–1.40)	NA
Ronco et al. (2007) Italy	25–60	ASC-US	22 466	22 708	1.17 (0.87–1.56)	0.84 (0.56–1.25)	0.58 (0.44–0.77)	0.62 (0.56–0.69)
Ronco et al. (2007) Italy	25–60	LSIL	22 466	22 708	1.03 (0.74–1.43)	0.72 (0.46–1.13)	0.58 (0.43–0.78)	NA
Strander et al. (2007) Sweden	23–60	ASC-US	8810	4674	1.40 (0.99–1.98)	NA	0.99 (0.74–1.33)	0.47 (0.27–0.82)
Maccallini et al. (2008) Italy	26–64	ASC-US	4299	4182	1.24 (0.72–2.15)	NA	1.40 (0.84–2.33)	0.31 (0.23–0.42)
Sykes et al. (2008) New Zealand	16–75	ASC-US	453	451	1.00 (0.83–1.21)	NA	NA	0.29 (0.16–0.55)
Siebers et al. (2008) Netherlands	30–60	ASC-US	40 047	48 941	1.00 (0.84–1.20)	1.05 (0.86–1.29)	0.99 (0.80–1.22)	NA
Siebers et al. (2009) Netherlands	30–60	ASC-US	40 047	48 941	NA	NA	1.03 (0.66–1.78)	NA
Klug et al. (2013) Germany	≥ 20	LSIL	9296	11 331	2.74 (1.66–4.53)	2.87 (1.55–5.32)	1.17 (0.81–1.67)	7.18 (2.55–20.2)

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; Conv., conventional cytology; LBC, liquid-based cytology; LSIL low-grade squamous intraepithelial lesion; NA, not available; PPV, positive predictive value; RR, relative risk.

[a] RR and 95% CI are reported as computed by the authors in main analyses, including adjustment procedures.

[b] Authors did not report relative measures or 95% CI; these have been computed from raw data by the Working Group.

compared with conventional cytology for both CIN2+ and CIN3+. The PPV was similar in the two groups. The authors only reported detection for LSIL+, and most of the abnormal cytology results, particularly for conventional cytology, were ASC-US. Surprisingly, the authors reported almost no unsatisfactory samples using conventional cytology. [The Working Group noted a high risk of bias in this study, and concern about generalizability.]

In RCTs, as well as in paired studies with unbiased assessment, LBC had a slightly higher sensitivity for detection of CIN2+ compared with conventional cytology; the difference in sensitivity, if any, in some studies seemed to be smaller for detection of CIN3+. This result in relation to sensitivity is consistent with that obtained in two very large population-based split-sample studies (Davey et al., 2007; Halford et al., 2010), which were not included in the systematic reviews on accuracy reported in the previous paragraph, because of incomplete assessment. In contrast, conventional cytology in most contexts had higher specificity for correctly classifying CIN1 or less severe conditions as negative, particularly when ASC-US was the test threshold, whereas the difference was smaller when LSIL was the test threshold. Because of this, the PPV was lower for LBC in many studies. Finally, a reduction in the proportion of unsatisfactory slides using LBC was reported in all studies, except for the study by Klug et al. (2013).

There is heterogeneity between studies, as is expected when comparing two tests that require expertise and training and for which not all countries use the same classification system. As reported before, the specificity of cytological tests at the threshold of ASC-US or LSIL is influenced by the prevalence of HPV infection in the tested population; this may explain part of the heterogeneity (Davey et al., 2006).

(c) Evidence on the effect of LBC on screening performance

Results from population-based studies have not always confirmed the data from randomized and paired-sample cross-sectional diagnostic accuracy studies. A summary of the characteristics of studies evaluating the effect that the introduction of LBC has had on screening performance and its effectiveness is reported in Table 4.16, and Table 4.17 summarizes the main results of comparisons between the performance of LBC and that of conventional cytology.

In England, Blanks & Kelly (2010) used aggregated routine quality assurance data from screening laboratories and reported an increase in PPV and a reduction in variability between laboratories after the introduction of LBC. Although differences observed in before-and-after studies may be due to other factors that changed concomitantly, this study compared a large number of laboratories and also investigated an outcome (variability between laboratories) that is directly linked to the introduction of the technology, but which should not be linked to trends in epidemiology or in differences in the screened population, making the causal link more plausible. [The Working Group noted adequate methodology in this study.]

In Iceland, Sigurdsson (2013) compared the results of LBC with conventional cytology in 2007–2011, when the organized screening programme shifted to LBC and other laboratories still used conventional cytology. The authors found no increase in the detection of CIN2+ or of CIN3+ in women younger than 40 years. In women older than 40 years there was a small, non-significant decrease in CIN3+ detection, whereas CIN2+ detection was similar in conventional cytology and LBC. The PPV of LBC was similar to or slightly higher than that of conventional cytology. The authors tried to adjust for differences observed between the results of the organized screening laboratory and the other

laboratories before the introduction of LBC. Nevertheless, the study design cannot exclude that observed differences between the performance of LBC and that of conventional cytology could be due to differences in the underlying populations and in the proficiency of the cytologists reading the slides. [The Working Group noted a very high risk of bias in this study.]

Gradual implementation of LBC in Japan apparently led to a 2-fold higher detection rate of CIN2+ and CIN3+. (Akamatsu et al., 2012). However, the analysis did not take into account differences in age, previous history of screening, and calendar time, i.e. those factors that could influence detection, with only raw numbers of tests performed and lesions found reported. [The Working Group noted a very high risk of bias in this study.]

A comparison before and after implementation of LBC with computer-assisted reading in Denmark (Rebolj et al., 2015) found slightly different results for the FocalPoint/SurePath system compared with the ThinPrep Imaging System. In this analysis, the effect of the introduction of LBC cannot be distinguished from the effect of the introduction of computer-assisted technology. Although ThinPrep had similar detection rates compared with conventional cytology, SurePath identified more CIN2+ and CIN3+. However, PPV was improved by 50% with ThinPrep but was 14% lower with the SurePath system than with conventional cytology. Neighbouring areas that continued using conventional cytology throughout the study period showed no changes, suggesting that any changes observed in the areas where LBC with computer-assisted cytology had been introduced were due to the new technologies. [The Working Group noted a high risk of bias in this study.]

One of the largest published studies comparing LBC with conventional cytology used data from the national screening programme in the Netherlands. Rozemeijer et al. (2016) reported an adjusted relative recall, compared with conventional cytology, that was slightly lower for ThinPrep and slightly higher for SurePath. The detection of CIN2+ was almost identical for ThinPrep and conventional Pap testing, and it was slightly higher with SurePath, with no significant difference in PPV between the three tests. Because the study included more than 3 million conventional Pap tests, 1.6 million ThinPrep slides, and 1.3 million SurePath slides, it had power to give very precise estimates adjusted for age, socioeconomic status, region, and calendar time. Furthermore, the national screening programme in the Netherlands started in 1980s, but the study covered the period 2000–2011; thus, even if the conventional Pap test was mostly used until 2005, there is no risk that the first rounds of screening, when detection is expected to be much higher, could bias the results. [The Working Group noted a low risk of bias in this study.]

Finally, the most recent population-based evaluation compared conventional Pap testing with LBC (a mix of 3 million ThinPrep slides and 757 320 SurePath slides) in opportunistic screening and organized screening in Japan (Ito et al., 2020). The referral rate was higher with LBC, as was the detection of CIN2+, but the detection of CIN3+ was similar. The PPV of LBC for detection of CIN2+ was slightly higher than that of conventional cytology, whereas the PPVs for detection of CIN3+ were almost identical. Relative estimates were adjusted for age, calendar period, and region. [The Working Group noted a low risk of bias in this study.]

In conclusion, results about sensitivity from these large population-based studies are quite consistent with those of the RCTs and paired-sample studies assessing cross-sectional test accuracy, but data on lower specificity or PPVs have not been confirmed in all programmes. The difference between early studies and these large population-based comparisons may depend on a learning curve for LBC. Indeed, most of the

Table 4.16 Characteristics of observational studies to assess the effect of the introduction of liquid-based cytology on screening performance and effectiveness

Reference Country	No. of women	Study design	Setting	Age at entry (years)	LBC procedure	Type of comparison	Reported end-points	Long-term outcomes
Blanks & Kelly (2010) England	~2.5 million 102 laboratories, 13 643 abnormal tests	Before-and-after analysis of aggregated quality assurance data from screening laboratories	Organized screening	25–64	ThinPrep; SurePath	Before and after in laboratories that shifted from Conv. to LBC during 2005–2008	PPV	No
Akamatsu et al. (2012) Japan	LBC: 29 119 Conv.: 49 108	Results for 2 consecutive rounds of screening during the shift from conventional Pap testing to LBC	Organized screening	NR	SurePath	Round 1: LBC vs Conv. Round 2: Conv. then Conv. Conv. then LBC LBC then LBC	Detection at round 1 and at round 2	Yes; CIN2+, CIN3+, and cervical cancer detection at next round
Sigurdsson (2013) Iceland	42 654 LBC tests in 20 439 women 103 909 Pap tests in 61 574 women	Comparison of conventional and LBC results in 2007–2011. Data adjusted for differences in pre-existing clinics before the introduction of LBC (2000–2004)	Spontaneous and organized screening	20–69	ThinPrep	LBC-observed outcomes vs expected outcomes	Relative detection Relative referral Relative PPV	No
Rebolj et al. (2015) Denmark	Conv. always: before, 47 300; after, 53 979 Conv. then SurePath: before, 23 849; after, 62 644 Conv. then ThinPrep: before, 33 614; after, 74 522	Before-and-after study with concomitant control	Organized screening	23–59	ThinPrep + ThinPrep Imaging System SurePath + FocalPoint + HPV triage for ASC-US	Conv. vs ThinPrep Conv. vs SurePath Before and after in areas that shifted from Conv. manual reading with repeat cytology for ASC-US to LBC + computer-assisted reading 1 area did not change during the study period	Relative referral Relative detection Relative PPV	No

Table 4.16 (continued)

Reference Country	No. of women	Study design	Setting	Age at entry (years)	LBC procedure	Type of comparison	Reported end-points	Long-term outcomes
Rozemeijer et al. (2016, 2017) Netherlands	Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817	Concomitant comparison in cohort study; women may change the exposure over time	Organized screening	29–63	ThinPrep; SurePath	Conv. vs ThinPrep Conv. vs SurePath SurePath vs ThinPrep Comparison of baseline outcomes (relative detection and relative referral) Long-term outcome: incidence of cancers after negative screening test	Relative referral Relative detection Cumulative detection of cancers after negative test	Yes; cumulative incidence of cervical cancer
Ito et al. (2020) Japan	3 815 131 ThinPrep: 3 057 810 SurePath: 757 321	Concomitant comparison in cohort study; women may change the exposure over time	Spontaneous and organized screening	≥ 20	ThinPrep; SurePath	Conv. vs any LBC Poisson regression to compare adjusted detection of CIN2+ and CIN3+	Relative detection Relative referral Relative PPV	No

ASC-US, atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; Conv., conventional cytology; LBC, liquid-based cytology; NR, not reported; PPV, positive predictive value.

Table 4.17 Results from observational studies on screening performance with liquid-based cytology compared with conventional cytology

Reference Country	No. of women	Referral for further assessment	Detection of CIN2+ and CIN3+	PPV
Blanks & Kelly (2010) England	~2.5 million 102 laboratories, 13 643 abnormal tests SurePath and ThinPrep	NA	NA	PPV for CIN3+: Before (Conv.): severe dysplasia, 75%; moderate, 37%; mild, 7% After (LBC): severe dysplasia, 79%; moderate, 37%; mild, 7% PPV for CIN2+: Before (Conv.): severe dysplasia, 88%; moderate, 70%; mild, 23% After (LBC): severe dysplasia, 90%; moderate, 72%; mild, 19%
Akamatsu et al. (2012) Japan	Conv.: 49 108 LBC: 29 119 SurePath and ThinPrep	NA	Conv.: CIN2+ (n = 123), 2.5/1000; CIN3+ (n = 66), 1.3/1000; cancer (n = 5), 0.10/1000 LBC: CIN2+ (n = 167), 5.7/1000; CIN3+ (n = 110), 3.8/1000; cancer (n = 13), 0.45/1000 RR LBC vs Conv.: CIN2+: 2.3 (95% CI, 1.8–2.9) CIN3+: 2.8 (95% CI, 2.1–3.9) Cancer: 4.4 (95% CI, 1.5–15.7)	NA
Sigurdsson (2013) Iceland	103 909 Pap tests in 61 574 women 42 654 LBC tests in 20 439 women	Observed/expected ratio for ASC-US+ cytology with LBC (expected computed according to cytology distribution before introduction of LBC): Women aged 20–39 yr: 1.27 (P that ratio is different from 1 < 0.001) Women aged 40–69 yr: 0.88 (P that ratio is different from 1 = 0.026)	Observed/expected ratio for CIN2+ with LBC (expected computed according to results before introduction of LBC): Women aged 20–39 yr: Observed/expected CIN2+: 1.06 (P that ratio is different from 1 = 0.36) Observed/expected CIN3+: 0.96 (P that ratio is different from 1 = 0.67) Women aged 40–69 yr: Observed/expected CIN2+: 0.75 (P that ratio is different from 1 = 0.82) Observed/expected CIN3+: 0.74 (P that ratio is different from 1 = 0.13)	PPV of ASC-US+ cytology for CIN2+: Women aged 20–39 yr: Conv.: 34.1% LBC: 34.8% Women aged 20–39 yr: Conv.: 16.1% LBC: 19.0%

Table 4.17 (continued)

Reference Country	No. of women	Referral for further assessment	Detection of CIN2+ and CIN3+	PPV
Rebolj et al. (2015) Denmark	Conv. always: before, 47 300; after, 53 979 Conv. then SurePath: before, 23 849; after, 62 644 Conv. then ThinPrep: before, 33 614; after: 74 522	Relative proportion of ASC-US+: Conv. always: 0.98 (95% CI, 0.91–1.07) SurePath vs Conv.: 1.99 (95% CI, 1.87–2.11) ThinPrep vs Conv.: 0.70 (95% CI, 0.66–0.75)	Relative before/after detection of CIN2+: Conv. always: 1.02 (95% CI, 0.88–1.18) SurePath vs Conv.: 1.71 (95% CI, 1.53–1.91) ThinPrep vs Conv.: 1.06 (95% CI, 0.93–1.21) Relative before/after detection of CIN3+: Conv. always: 1.10 (95% CI, 0.93–1.30) SurePath vs Conv.: 1.66 (95% CI, 1.46–1.88) ThinPrep vs Conv.: 0.99 (95% CI, 0.85–1.15)	Relative before/after PPV of ASC-US cytology for CIN2+: Conv. always: 1.03 (95% CI, 0.92–1.16) SurePath vs Conv.: 0.86 (95% CI, 0.79–0.94) ThinPrep vs Conv.: 1.51 (95% CI, 1.36–1.68) Relative before/after PPV of ASC-US cytology for CIN3+: Conv. always: 1.12 (95% CI, 0.97–1.29) SurePath vs Conv.: 0.83 (95% CI, 0.75–0.93) ThinPrep vs Conv.: 1.41 (95% CI, 1.23–1.61)
Rozemeijer et al. (2016, 2017) Denmark	Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817	OR of cytology ≥ borderline or mild dyskaryosis: ThinPrep vs Conv.: 0.96 (95% CI, 0.93–0.99) SurePath vs Conv.: 1.12 (95% CI, 1.09–1.16)	OR of cytology having a CIN2+ detected: ThinPrep vs Conv.: 0.99 (95% CI, 0.96–1.02) SurePath vs Conv.: 1.08 (95% CI, 1.05–1.12)	OR: PPV of cytology ≥ borderline or mild dyskaryosis for histology: ThinPrep vs Conv.: CIN2: 1.08 (95% CI, 0.99–1.17) CIN3: 1.06 (95% CI, 0.99–1.13) Cancer: 0.98 (95% CI, 0.83–1.15) SurePath vs Conv.: CIN2: 1.06 (95% CI, 0.98–1.15) CIN3: 0.97 (95% CI, 0.91–1.03) Cancer: 0.94 (95% CI, 0.80–1.10)
Ito et al. (2020) Japan	ThinPrep: 3 057 810 SurePath: 757 321	Conv.: 1.13% (34 435) LBC: 1.49% (11 443) Crude RR, 1.32 (95% CI, 1.30–1.35)	Adjusted RR, LBC vs Conv.: CIN2+: 1.16 (95% CI, 1.08–1.25) CIN3+: 1.00 (95% CI, 0.90–1.11)	Adjusted RR, LBC vs Conv.: CIN2+: 1.17 (95% CI, 1.09–1.26) CIN3+: 1.01 (95% CI, 0.91–1.12)

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; LBC, liquid-based cytology; NA, not applicable; OR, odds ratio; PPV, positive predictive value; RR, relative risk; yr, year or years.

initial studies were conducted by cytologists whose university education had been based on conventional cytology and who were retrained to LBC, whereas these large population-based studies also included cytologists who have had more experience with LBC in their routine work; some cytologists even started their professional activity using LBC. It is not possible to determine whether the new generation of cytologists, who began their studies and training with LBC in the USA and, more recently, in many countries in Europe and Asia, would produce different values of relative sensitivity and, particularly, of specificity.

(d) Evidence on the effectiveness of LBC in routine cervical screening programmes

The aim of cervical cancer screening is to prevent cancer incidence through the detection and treatment of CIN2+ lesions. However, there is evidence that only 30% of CIN3 lesions progress to cancer in a 30-year time span (McCredie et al., 2008), and this proportion is even lower for CIN2 lesions. Most CIN2 lesions, and also CIN3 lesions, will regress spontaneously (Ronco et al., 2008) or persist without progression. Therefore, an increase in CIN2+ detection is an advantage only if it includes those lesions that would progress to cancer or at least would persist for a long time. To test the efficacy of LBC with this longitudinal approach, it is necessary to conduct studies with a long-term follow-up of women who tested negative in one of the two tests and to observe the cumulative incidence of cancer or CIN3 as a surrogate of cancer risk. This is not possible with paired-sample studies, because in these studies women are managed (i.e. assessed and eventually treated) according to the results of both tests. Only RCTs with long-term outcome assessment and concurrent cohort studies can provide a longitudinal approach.

One RCT (Strander et al., 2007) and two observational studies (Akamatsu et al., 2012; Rozemeijer et al., 2017) have published results comparing the cumulative incidence of CIN or cervical cancer after a negative test result from LBC or conventional cytology screening (Table 4.18).

The RCT in Sweden (Strander et al., 2007) reported a cumulative incidence from 1.5 years after recruitment up to 3 years and 7 months (i.e. excluding lesions found at recruitment, but including those found at the next screening round) of 6 per 1000 for LBC and 5.3 per 1000 for conventional cytology (RR, 1.12; 95% CI, 0.68–1.83). [The Working Group noted a low risk of bias and a very imprecise estimate in this study.]

Akamatsu et al. (2012), in Japan, reported a lower detection of CIN2+, CIN3+, and invasive cancer after LBC (SurePath); the numbers were small, however, and the difference may have been due to chance. Furthermore, the populations screened with LBC and conventional cytology were not comparable, but the authors could not adjust for possible confounders. [The Working Group noted a very high risk of bias in this study.]

Finally, the largest study compared the cumulative incidence of invasive cancer after conventional cytology and two different LBC systems (ThinPrep and SurePath) in the national screening programme in the Netherlands (Rozemeijer et al., 2017). The authors adjusted the estimates for age, socioeconomic status, calendar period, and region and found very similar incidence rates of cancer detected by LBC and conventional cytology (Table 4.18); SurePath showed a significant reduction in cancer incidence compared with both conventional cytology and ThinPrep. [The Working Group noted a low risk of bias in this study.] A previous study in the Netherlands comparing two smaller cohorts from the national screening programme, one screened with conventional cytology and one with LBC, found a 50% lower occurrence of CIN2+ in a follow-up of about 1.5 years after a negative LBC test result compared with conventional cytology (7 of 34 219 vs 21 of 49 856; $P = 0.091$) (Beerman

et al., 2009); some of the women included in this study may be also included in the study of Rozemeijer et al. (2017).

4.3.4 Cytology based on Romanowsky–Giemsa staining

(a) Definition of Romanowsky–Giemsa staining

The term "Romanowsky–Giemsa staining" or "Romanowsky staining" refers to several techniques used to stain cytological specimens, in which the Romanowsky effect is used to differentiate the cell components through different colour hues (Theil, 2012; Bezrukov, 2017), in particular the purple staining of chromatin. Nuclei stained with these techniques show variations in staining that enable characterization of their morphology. The technique is named after Romanowsky (Krafts & Pambuccian, 2011). The effect is based on the use of two dyes, eosin and a methylene blue that has been subject to oxidative demethylation. This dye, called polychrome methylene blue, is a mix of several molecules, including methylene blue, azure A, azure B, azure C, thionine, methylene violet Bernthsen, methyl thionoline, and thionoline (Marshall, 1978).

Techniques based on the Romanowsky effect have been used for a long time to stain many types of cytological specimens, and are still the standard for the diagnosis of infection with *Leishmania* and other disease-causing microorganisms, such as *Plasmodium* (malaria), *Toxoplasma*, and *Pneumocystis* (Marshall, 1978; Horobin, 2011; Li et al., 2012; Bain, 2017). The technique is also still used to stain haematological smears (Horobin, 2011; Theil, 2012; Bain, 2017).

For gynaecological cytology, the technique has been completely replaced by Pap staining (Spriggs, 1977; Broder, 1992; Solomon et al., 2002) except for in some countries of the former Soviet Union.

(b) Differences between Romanowsky–Giemsa staining and Pap staining

Romanowsky–Giemsa staining was developed for air-dried specimens, whereas the Pap stain is used for wet-fixed specimens. Wet fixation enables better differentiation of nuclear chromatin structures, particularly nucleoli, and better characterization of nuclear shape abnormalities that are present in neoplastic cells (Krafts & Pambuccian, 2011). Another limitation of the Romanowsky–Giemsa stain compared with the Pap stain is its inability to characterize cytoplasmic keratinization, a feature that is particularly important in the diagnosis of squamous cell neoplasia (Krafts & Pambuccian, 2011). Finally, the Romanowsky–Giemsa stain does not penetrate well into the small, three-dimensional groups of cells that may be present in cytological specimens; this results in an absence of staining in inner cells. In contrast, the Pap stain method can stain small groups of overlapping cells (Krafts & Pambuccian, 2011).

The Romanowsky–Giemsa stain also has advantages. For example, in air-dried specimens the differences between the nuclear and cytoplasmic diameters are magnified, which is useful in distinguishing potential cellular transformation (Boon & Tabbers-Bouwmeester, 1980; Boon & Drijver, 1986). Chromatin is hyperchromatic, which enables a better impression at low magnification, but there is reduced detail of the nuclear structures at higher magnifications. Some cytoplasmic structures are better defined, and chondroid cytoplasmic material can be identified (Krafts & Pambuccian, 2011). Also, a Leishman–Giemsa cocktail, which is based on two staining solutions, both of which produce the Romanowsky effect, enables better staining of nuclei, on the basis of chromatin, vesicularity, and membrane integrity, and higher quality of cytoplasm staining, on the basis of the transparency and nature of the cell membrane, compared with Pap staining (Padma et al., 2018).

Table 4.18 Long-term outcomes of cervical cancer screening by liquid-based cytology compared with conventional cytology

Reference Country	Design	No. of women	Detection	IRR or RR
Strander et al. (2007) Sweden	RCT with 3 yr and 7 mo follow-up	Conv.: 8810 LBC: 4674	CIN2+ detection during follow-up from 1.5 yr to 3 yr and 7 mo after recruitment, all screened as routine: After LBC: 0.60% (28/4674) After Conv.: 0.53% (47/8810)	RR, 1.12 (95% CI, 0.68–1.83)
Akamatsu et al. (2012) Japan	Results for 2 consecutive rounds of screening during the shift from conventional Pap testing to LBC	Conv. then Conv.: 73 253 Conv. then LBC: 33 318 LBC then LBC: 51 723	Conv. then Conv.: CIN2+ (n = 115), 1.6/1000; CIN3+ (n = 58), 0.8/1000; cancer (n = 10), 0.14/1000 Conv. then LBC: CIN2+ (n = 38), 1.1/1000; CIN3+ (n = 24), 0.7/1000; cancer (n = 2), 0.06/1000 LBC then LBC: CIN2+ (n = 41), 0.8/1000; CIN3+ (n = 24), 0.5/1000; cancer (n = 1), 0.02/1000	LBC then LBC vs Conv. then LBC: CIN2+: RR, 0.70 (95% CI, 0.44–1.11) CIN3+: RR, 0.64 (95% CI, 035–1.18) Cancer: RR, 0.32 (95% CI, 0.01–6.19)
Rozemeijer et al. (2017) Netherlands	Concomitant comparison in cohort study; women may change the exposure over time	Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817	72 mo cumulative incidence of cervical cancer after normal cytology: Conv.: 1042 cancers; 13 796 018 person-yr ThinPrep: 328 cancers; 5 201 188 person-yr SurePath: 231 cancers; 4 835 917 person-yr	Adjusted IRR, SurePath vs Conv., 0.81 (95% CI, 0.66–0.99) Adjusted IRR, ThinPrep vs Conv., 1.15 (95% CI, 0.95–1.38)

CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; Conv., conventional cytology; IRR, incidence rate ratio; LBC, liquid-based cytology; mo, month or months; NA, not applicable; PPV, positive predictive value; RCT, randomized controlled trial; RR, relative risk; yr, year or years.

Finally, the main advantage of the Romanowsky–Giemsa stain is that the procedure for preparing the slides is less time-consuming and uses reagents that are less expensive and easier to obtain (Jarynowski, 2019).

(c) Use of the technology

Romanowsky–Giemsa staining is used for gynaecological cytology in countries of the former Soviet Union, where it is described mostly with the following names: Romanowsky–Giemsa, May–Grünwald–Giemsa, and Pappenheim (Rogovskaya et al., 2013).

The first official document describing the application of the Romanowsky–Giemsa stain for cervical specimens was published in 1976 when Order No. 1253 was issued by the Ministry of Health of the Soviet Union. With almost no changes, the method was used until the dissolution of the Soviet Union and the emergence of the newly independent states. Table 4.19 lists documents stating recommendations for the use of Romanowsky–Giemsa staining in cervical cancer screening in the countries of the former Soviet Union.

There are few reports on the change in cervical cancer screening methods in countries of the former Soviet Union. Three Baltic countries – Estonia, Latvia, and Lithuania – became part of the European Union and implemented Pap-based cervical cancer screening programmes in 2004–2006. In Belarus, Pappenheim staining (a modification of Romanowsky–Giemsa staining) is used (IARC, 2012). In Kazakhstan, services successfully moved to Pap-based screening in 2008 (Aimagambetova & Azizan, 2018; Bekmukhambetov et al., 2018). In the Republic of Moldova, a shift to Pap testing started after 2016, but barriers related to cost and training have been described (Davies et al., 2016; Jarynowski, 2019). Analysis of cervical screening services in the Republic of Moldova by an external adviser for the ministry of health also pointed out that the absence of an international community for standardization makes quality improvement difficult (Davies et al., 2016).

In the Russian Federation, where cervical screening is budgeted by region, some countries have changed to Pap testing. Since 2019, the Ministry of Health of the Russian Federation has recommended against the use of Romanowsky–Giemsa staining for cervical screening (see Table 4.19). Implementation of this recommendation was affected by several barriers, including the higher costs of the reagents and the need for complete retraining of cytotechnicians and cytologists. In Ukraine, there is no clear document recommending a shift from cytology based on Romanowsky–Giemsa staining to Pap testing, mostly because of economic barriers to the implementation of Pap testing.

In other countries in central Asia, the situation is unclear. In 2017, the United Nations Population Fund (UNFPA) funded a project on the use of VIA in Tajikistan (UNFPA, 2019), which suggested that infrastructure for cytology was not sufficient. In Turkmenistan, Pap staining followed by retesting with Romanowsky–Giemsa, or HPV testing, is replacing the use of cytology based on Romanowsky–Giemsa staining as a stand-alone technique because of an improvement in economic resources compared with other countries in central Asia; however, the coverage is probably low (Rogovskaya et al., 2013).

(d) Epidemiology of cervical cancer in countries in eastern Europe and central Asia

WHO data on cervical cancer mortality from 1975 to 2005 show a different trend in most eastern European countries compared with western European countries (La Vecchia et al., 2010). In general, most western European countries had a decreasing trend, whereas in eastern European countries mortality rates were essentially stable or had a slight increasing trend (see also Section 1.1.1, Fig. 1.5), with the exception of

Table 4.19 Former and current use of cytology based on Romanowsky-Giemsa staining

Country	Position of official guidelines	Use of technology	References
Belarus	NA	Pappenheim staining observed during the IARC visit to the National Cancer Centre in Minsk in February 2019 for the IARC-WHO Regional Office for Europe training course	IARC (2012)
Kazakhstan	NA	Mainly opportunistic screening by cytology based on Romanowsky–Giemsa staining until 2007. From 2008, 60% of all smears were prepared using the Pap stain and 40% using the Romanowsky–Giemsa stain. Since 2009, 100% of screening smears use Pap staining	Aimagambetova & Azizan (2018)
Republic of Moldova	Recommendation to progressively change from Romanowsky–Giemsa staining to Pap staining during the course of 2017	Opportunistic screening, with the majority using Romanowsky–Giemsa staining	Davies et al. (2016)
Russian Federation	Order No. 124 (13 March 2019): Pap testing only. This effectively cancelled the previous Order No. 869 (26 October 2017), when Romanowsky–Giemsa staining was officially mentioned	Cervical smear test with Romanowsky–Giemsa or May–Grünwald–Giemsa staining. Until 2017, annual examinations for women aged ≥ 18 yr or after first intercourse, with no upper age limit. Moscow used a screening age range of 35–69 yr and a screening interval of 3 yr Officially, it should now be Pap testing only. However, some centres still use Romanowsky–Giemsa staining because it is less expensive. [The regions are responsible for budgets.]	Olson et al. (2016); Ministry of Health of the Russian Federation (2019)
Soviet Union	Order No. 1253 (30 December 1976) introduced the use of Romanowsky–Giemsa staining across the whole country	In 1964, annual cytology screening was introduced in the former Soviet Union as part of routine cervical cancer screening; in 1976, the Ministry of Health of the Soviet Union established centralized cytology laboratories in all regions and republics. Opportunistic basis, using Romanowsky–Giemsa staining or haematoxylin and eosin staining	Rogovskaya et al. (2013); Olson et al. (2016)
Turkmenistan	Order No. 144 (2014)	Pap testing followed by Romanowsky–Giemsa staining or HPV testing	WHO (2019)

HPV, human papillomavirus; NA, not available; yr, year or years.

Czechia, where data are available only since 1987 and there has been a slight decrease.

More detailed analyses from the Russian Federation (Barchuk et al., 2018) showed a slight decline in cervical cancer mortality rates until the early 1990s, followed by a slight increase after the mid-1990s. An increasing trend in incidence rates has been observed since 1989, when data are available for the Russian Federation (Barchuk et al., 2018), and from other countries in eastern Europe and central Asia (Bruni et al., 2019a, b). A detailed analysis from the Arkhangelsk Regional Cancer Registry in the north-west of the Russian Federation (Grjibovski et al., 2018) found that both incidence and mortality rates increased from 2000 to 2014 but incidence increased more than mortality, showing that survival for women with cervical cancer had improved, possibly because of earlier detection and management; consistent with this, incident cancers showed a simultaneous shift to earlier stages at diagnosis. These figures suggest two conflicting effects: an increased risk of occurrence and an improvement in the early diagnosis of cancers. It is impossible to tell whether this improvement in early diagnosis also affected incidence through detection and treatment of precancerous lesions, but if an effect is present it is not sufficiently strong to reverse the increase in risk, probably as a result of an increase in HPV prevalence.

Countries in eastern Europe and central Asia have the highest incidence of cervical cancer in Europe, independent of the screening test coverage that they reported (Ferlay et al., 2018; Bruni et al., 2019a, b; Arbyn et al., 2020).

(e) Evidence on accuracy and effectiveness

(i) Accuracy

Limited data were found comparing the diagnostic accuracy of Romanowsky–Giemsa staining and the Pap test. Romanowsky–Giemsa staining (90%) has a lower specificity than the Pap test (98%) to distinguish cervical precancer; this is usually mitigated by repeating the test to reduce the possibility of missing women with precancer (Davies et al., 2016; Jarynowski, 2019). No data were available on sensitivity and on how it may be affected by repeating tests to increase specificity.

(ii) Performance in screening programmes

Table 4.20 summarizes the data on the performance of Romanowsky–Giemsa staining in screening programmes that are available in the peer-reviewed and grey literature.

Data on performance can provide some information about or insights into the accuracy of the programme, taking into consideration that the first-level test is usually the main determinant of programme accuracy, but it is not the only one. Furthermore, performance indicators are strongly influenced by the quality of routinely collected data, particularly the detection rate, because missing even a few lesions may lead to a large underestimation of the indicator. The few documents reporting the proportion of unsatisfactory samples when using Romanowsky–Giemsa staining show values that are close to or higher than the upper bound of the range observed in western European countries with Pap staining, i.e. about 10%. When data from different laboratories enable benchmarking (Davies et al., 2016), the proportion of unsatisfactory slides varies widely (ranging from 0% to 5.7%), which suggests low reproducibility of the technique. [This heterogeneity may come from differences in the way in which cytologists from different laboratories interpret the findings or from variability in how samples are collected and processed.] A high variability between laboratories in the detection rates of LSIL and HSIL was also reported.

The detection rate varies widely. [It is not clear whether the available data report histologically confirmed cases or simply the cytological classification (which would be not a detection rate but a proportional analysis of positives).] In

some cases (Iskhakova et al., 2012; Table 4.20), the detection rate is very low compared with the cervical cancer incidence in the region; [this suggests that the programme has poor sensitivity or that there is underreporting of histological findings].

Data on referral rates were not identified. [It is not clear how women with abnormal findings are managed, i.e. whether with direct referral for colposcopy or with repeated cytology.] Consequently, no data on PPV were found or could be estimated from the available reports.

(iii) Efficacy and effectiveness

No trials have been identified that compare the efficacy of cytology based on Romanowsky–Giemsa staining with that of Pap testing or other cytological staining techniques.

No controlled studies on the effect of screening programmes on cervical cancer incidence or mortality have been identified.

Time-trend studies conducted after the dissolution of the Soviet Union in 1989, as well as data from routine cancer statistics, showed no reduction and in some cases an increase in cervical cancer incidence and mortality rates in most of the countries where Romanowsky–Giemsa staining is used for screening (La Vecchia et al., 2010; Barchuk et al., 2018; Grjibovski et al., 2018; Bruni et al., 2019a, b). This trend is common to almost all countries in eastern Europe and central Asia, except Czechia, independent of the method of cytology staining used and of the reported coverage of the screening test (La Vecchia et al., 2010; Bruni et al., 2019a, b).

4.3.5 Harms of cytological techniques

(a) Physical harms

Pelvic examination is a very sensitive medical procedure, and special considerations are needed. Bloomfield et al. (2014) performed a systematic review of pelvic examination in asymptomatic, non-pregnant, average-risk adult women. Eight studies including 4576 women reported that women experienced pain or discomfort; the median rate was 35%, and rates ranged from 11% to 60%. Rates of fear, embarrassment, or anxiety ranged from 10% to 80%. Pain can be exacerbated by atrophic vaginal mucosa and vaginal dryness in menopausal women (Elit, 2014). However, some studies conducted in the United Kingdom reported that younger women experience more embarrassment and pain than older women (Yu & Rymer, 1998; Fiddes et al., 2003).

Although female patients usually prefer a female physician for gynaecological examinations, one study in 167 women with median age 25 years in the USA found that pain scores for examinations by male physicians and female physicians were not significantly different (Moettus et al., 1999).

In a cross-sectional study reporting on the pain and physical discomfort experienced during a Pap test, Hoyo et al. (2005) carried out a questionnaire survey of 144 African American women aged 45–65 years. They reported that 45.8% of women who did not attend screening and 17.5% of women who attended screening experienced pain during the cytological examination ($P < 0.0001$). Women who felt pain during the cytological examination were less likely to participate in further cervical cancer screening. In a study in Vietnamese American women aged 18–64 years, 55% of 240 women who had had cytology within 3 years reported that concern about pain or discomfort was a barrier to cytological examination (OR, 0.5; 95% CI, 0.3–1.1) (Taylor et al., 2004). In a longitudinal cohort study in 490 sexually active young women aged 12–24 years who presented to a hospital-based adolescent clinic in the USA, Kahn et al. (2003) reported that women who returned for a follow-up visit were more likely to believe that the follow-up Pap test would not be painful compared with those who did not return (77% vs 65%, OR, 1.73; 95% CI, 1.08–2.83).

Table 4.20 Performance and detection rate of precancerous lesions using Romanowsky–Giemsa staining

Reference	Country Date	Setting	Population/no. of tests (*N*)	Unsatisfactory samples	Detection rate (%)
Iskhakova et al. (2012)	Russian Federation 2009–2011	Meleuz (Bashkortostan), centralized cytological laboratory	79 710 women, aged 20–60 yr	Unsatisfactory: 0.6% Insufficiently satisfactory: 15.6% Satisfactory: 80.3%	CIN1: 0.3 (*n* = 168) CIN2: 0.2 (*n* = 86) CIN3: 0.05 (*n* = 43) CIS: 0.02 (*n* = 13) Cervical cancer: 0.02 (*n* = 13)
Kozyreva et al. (2012)	Russian Federation	Vladikavkaz (North Ossetia), oncological dispensary	9525 nuclei of malignant and normal cervical cells	NR	Number detected (%): CIN1: 530 (0.056%) CIN2: 960 (0.100%) CIN3: 890 (0.093%)
Chernyakova (2016)	Ukraine 2015	Kharkiv, university clinic	37 women aged 20–64 yr enrolled in "opportunistic screening"	19 women – inflammation; *Chlamydia* in 9.6%, HPV in 28.5% CIN1 in 2/37 Cytology–colposcopy discrepancy in 5/37	NR
Davies et al. (2016)	Republic of Moldova 2015	National audit Data from 7 of the largest laboratories	236 579 smears	Between laboratories, proportion of abnormal results varied from 0.32% to 6.06%, and unsatisfactory results varied from 0.0% to 5.7%	Range between laboratories: ASC-US: 0.04–0.64 LSIL: 0.02–2.35 HSIL: 0.02–2.10 AGUS: 0.0–0.01 ASC-H: 0.0–0.26 Cervical cancer: 0.0–0.18
Aktanko et al. (2018)	Russian Federation	Vladivostok	4032 women, aged > 25 yr	NR	CIN1: 21.9 (*n* = 20) CIN2: 12.1 (*n* = 11) CIN3: 19.7 (*n* = 18) CIS: 4.4 (*n* = 4) SCC: 30.7 (*n* = 28) Adenocarcinoma: 1.09 (*n* = 1)
Grebenkina et al. (2018)	Russian Federation 2018	Nizhny Novgorod, reference cytological centre; evaluated 10% of all cytological and 100% of all indeterminate samples	9415 cytological smears 12% processed by Romanowsky–Giemsa	23% of all slides (not only Romanowsky–Giemsa stained smears)	21–36% did not match the final diagnosis (including 2 missed cervical cancers)

Table 4.20 (continued)

Reference	Country Date	Setting	Population/no. of tests (*N*)	Unsatisfactory samples	Detection rate (%)
Kirillina et al. (2018)	Russian Federation 2017	Yakutia, different women's clinics	7600 women, aged 18–88 yr	Non-informative material: 1.9% Glandular epithelium not taken: 19.4%	All CIN+: 4.7 (*n* = 359) CIN1: 61.3 (*n* = 220) CIN2: 24.5 (*n* = 84) CIN3: 10.6 (*n* = 38) CIS + cervical cancer: 1.1 (*n* = 4, cervical cancer = 2)

AGUS, atypical glandular cells of undetermined significance; ASC-H, atypical squamous cells cannot exclude high-grade; ASC-US, atypical squamous cells of undetermined significance; CIN, cervical intraepithelial neoplasia; CIS, carcinoma in situ; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesions; LSIL, low-grade squamous intraepithelial lesions; NR, not reported; SCC, squamous cell carcinoma; yr, year or years.

When stirrups are used during the pelvic examination and cervical sampling, women are compelled to be in the dorsal lithotomy position, which can cause discomfort. Seehusen et al. (2006) measured physical and psychological effects in an RCT of 197 women who underwent gynaecological examinations in the USA with stirrups ($n = 97$) or without stirrups (women were examined with their feet placed on the corners of a fully deployed table extension; $n = 100$). All the women were draped with a full-sized sheet in a standardized manner that maximized the coverage of the body and enabled visualization of the perineum. Physical discomfort was higher in women who were examined with stirrups compared with those examined without stirrups (30.4% vs 17.2%). There was no significant reduction in sense of loss of control.

Korfage et al. (2012) sent questionnaires to the home addresses of 789 screening participants in the Netherlands before screening, after screening, and again with the screening results, to assess the effect of cervical cancer screening on health-related quality of life in women with normal test results. A female age-matched reference group ($n = 567$) was included. Although the average age was not significantly different between the groups (45.3 years vs 45.8 years; $P = 0.29$), the proportion of postmenopausal women was unknown. About 40% of screening participants experienced at least one of the following symptoms at least 1 day after the smear had been taken: lower abdominal pain, vaginal bleeding, discharge, urinary problems, or feeling sick. These symptoms were very painful or fairly painful for 12% of women.

(b) Psychological harms

Psychological harms can be experienced: (i) when samples are collected, (ii) as a result of waiting time to receive the results, (iii) from unsatisfactory smears, (iv) from abnormal results, and (v) upon follow-up because of abnormal results. All women in whom smears are taken have the potential to experience the first kind of harm. The potential effect of the second kind of harm will vary depending on the woman's previous knowledge and experience of cervical cancer screening. Other harms are limited to women with unsatisfactory and abnormal test results.

(i) All participants

Women naturally feel some personal embarrassment and discomfort when smears are taken for cervical cancer screening, as described above. In the Netherlands, Korfage et al. (2012) assessed the effect of cervical cancer screening on health-related quality of life in 789 women with normal test results, compared with a reference group ($n = 563$). Screening-specific anxiety was lower in the screened women than in women who had not been screened. When results before and after screening were compared, the EQ-5D rating of own health increased, the mental health score (Mental Component of the Short-Form 12) increased, and the general anxiety score (Spielberger State-Trait Anxiety Inventory [STAI-6]) decreased. There were no differences between the results in younger and older women. Although 19% reported a feeling of shame, pain, inconvenience, and nervousness during the smear-taking procedure, 80% of women were satisfied with their results after the cytology procedure.

In an interview survey in 13 women of various ages and backgrounds by Larsen et al. (1997), nearly all the women who had pelvic examination indicated that they were nervous before the consultation, but they regarded the examination as a necessary procedure for diagnosis. The women identified several factors that affected their ability to feel in control during the procedure, such as the physician's gender, informed communication, positioning during the examination, a feeling of lost integrity while naked, and trust in the physician. Yanikkerem et al. (2009) also emphasized the necessity of providing information during the gynaecological examination,

based on a questionnaire survey of 433 women who attended a gynaecological outpatient clinic in Turkey. In an interview survey of 262 women aged 21–65 years by Norrell et al. (2017), 62% of participants believed that open communication with their health-care provider was helpful in understanding the purpose and value of a pelvic examination. In further cohort studies in the USA and Europe, good communication positively affected the screening experience and improved screening adherence (Taylor et al., 2004; Thangarajah et al., 2016; Freijomil-Vázquez et al., 2019).

(ii) Women waiting to receive cytology results

Freijomil-Vázquez et al. (2019) carried out an interview survey of 21 women aged 21–52 years with confirmed diagnosis of CIN recruited from a gynaecology clinic in Spain. When health-care providers gave limited information about diagnosis, the women's anxiety increased as a result of the uncertainty and lack of decision-making ability they felt about the prevention and treatment of CIN.

In a questionnaire survey by Korfage et al. (2012), general anxiety and screen-specific anxiety levels were compared before and after Pap tests in 789 women in the Netherlands. A female age-matched reference group including 567 randomly selected women (aged 30–70 years) who were not due for cervical cancer screening within the next 2 years were sent a questionnaire through the regional screening organization in Maastricht. Screening participants reported less screen-specific anxiety ($P < 0.001$) than the reference group before screening, after screening, and also after the receipt of test results. After a normal result was received, general anxiety, as judged by the STAI-6, decreased slightly. Screen-specific anxiety measured using the Psychological Consequences Questionnaire increased initially but then decreased after the receipt of the Pap test results.

(iii) Women with unsatisfactory test results

The rates of unsatisfactory test results differ between screening programmes (see Section 4.3.1), and women with unsatisfactory test results can have higher levels of anxiety compared with women with normal test results. French et al. (2004) studied the psychological effects in 180 women with unsatisfactory smears and 226 women with normal results in the United Kingdom. Women with unsatisfactory test results had higher scores for state anxiety (STAI-6) and concern about the test results, perceived themselves to be at a higher risk of cervical cancer, and were less satisfied with the information they received about their test result compared with women with normal test results.

(iv) Women with abnormal test results

Most studies that reported the anxiety experienced by women after receiving an abnormal cytology test result were cross-sectional questionnaire surveys. Some studies investigated the duration of the psychological effects after receiving an abnormal test result.

Maissi et al. (2004) performed a questionnaire survey and compared the psychological effects in 366 women with normal results and 1010 women with abnormal test results (borderline or mild dyskaryosis) in the United Kingdom. Women with normal results had significantly lower scores for state anxiety (STAI-6), emotional distress (12-item General Health Questionnaire [GHQ-12]), and concern about test results than those with abnormal results. Similar findings were reported by Wardle et al. (1995), also in the United Kingdom.

A study in Sweden reported the results of a questionnaire survey of 242 women with two consecutive Pap tests reported as mild dysplasia (CIN1) who should, as a consequence, have undergone colposcopy and biopsy according to an agreed general programme (Ideström et al., 2003). Most women were satisfied with the follow-up; 72% felt they understood the meaning

and consequences of having mild dysplasia. Nevertheless, 59% reported feeling worried and anxious. Moreover, 30% of women thought that the results affected their daily life because of the stress induced by the need for additional testing, and 8% reported a negative influence on sexuality and their experience of sexual intercourse as a consequence of the management of mild dysplasia.

In the Trial of Management of Borderline and Other Low-grade Abnormal smears (TOMBOLA) study conducted in the United Kingdom, Gray et al. (2006) performed a questionnaire survey of the psychological and psychosocial effects in 3671 women with a low-grade abnormality (borderline nuclear abnormalities or mild dyskaryosis). On the Hospital Anxiety and Depression Scale (HADS), 57% of women had no anxiety (a score of < 8 is defined as a cut-off point for anxiety by the HADS anxiety subscale), 20% had scores consistent with some level of anxiety (scored 8–10), and 23% had scores that indicated a probable clinically significant level of anxiety (scored ≥ 11). Most women (91%) were classed as non-cases on the depression subscale (a score of < 8 is defined as a cut-off point for no depression by the HADS depression subscale). Statistically significant associations were found between reported anxiety and younger age, increased physical activity, ever having had a child, and current smoking status. There was also a strong association between anxiety and depression scores: 95% of women who scored ≥ 8 on the depression subscale also scored ≥ 8 on the anxiety subscale. In a multivariate analysis, significant associations were found between anxiety and worries about general health, feelings about self, worries about cervical cancer, future fertility, sex life, perceived risk of cervical cancer, and support received.

Pirotta et al. (2009) assessed the psychological effects of an abnormal Pap test result in 333 women aged 18–45 years in Australia who completed a survey 3 months after receiving their test results. General health-related quality of life scores were assessed using the EuroQol Visual Analogue Scale, in which participants select their current health status on a scale from 0 (death) to 100 (perfect health). The results were nearly equal in women with a normal smear, women with an abnormal smear, and women with confirmed CIN. The scores for worries and concerns, emotional impact, and control using the Human Papillomavirus Impact Profile were higher in women with abnormal Pap tests and CIN than in women with normal Pap tests. Concerns about effects on sex life and self-image were observed in women with high-grade lesions or external genital warts, but not in those with low-grade lesions.

Korfage et al. (2010) sent questionnaires to 270 women with borderline or mild dyskaryosis (BMD) test results in the previous 6–24 months identified through a regional screening organization, to evaluate general quality of life, general anxiety, and screen-specific anxiety. A similar questionnaire was sent to 372 randomly selected women (aged 30–60 years) who were due for screening (reference group). The women in the BMD group were younger than the women in the reference group (mean age, 43 years vs 46 years; $P < 0.001$); the proportion of postmenopausal women was unknown. Women in the BMD group had higher levels of general anxiety and screen-specific anxiety than those in the reference group; 44% of the BMD group had high anxiety (indicated by an STAI-6 score > 44) compared with 33% in the reference group ($P < 0.001$). This finding remained significant after adjustment for differences in age, job and marital status, having children or not, and country of birth. Although both groups reported positive attitudes towards the cervical cancer screening programme, women in the BMD group were more likely to report fear of cervical cancer as their reason for having a repeat smear taken, compared with women in the reference group (23% vs 4%; $P < 0.001$).

A questionnaire-based study in Germany to assess the psychological effect of an abnormal test result invited 595 women who had been referred to a special outpatient clinic with CIN for further evaluation (Thangarajah et al., 2016). Most of the women (68.8%) reported that they felt anxious on receipt of the test result, 26.3% felt panic, and 18.6% did not understand what the test result meant. After speaking with their physicians, 54.4% of women remained worried, 24.4% felt reassured, and 20.2% felt confident.

In an RCT in Norway, women were randomized to either hrHPV testing every 5 years (followed by cytology if hrHPV-positive; $n = 487$) or cytology testing every 3 years (followed by hrHPV testing if low-grade cytology was detected; $n = 521$); anxiety and depression scores were compared by screening group and by test result (Andreassen et al., 2019). The mean age was 51 years and was similar in both study groups. The frequency of abnormal primary cytology results (≥ ASC-US) was 54% and of positive primary hrHPV test results was 53%. Compared with women who were screened with cytology, women screening with hrHPV were not more likely to experience mild anxiety and depression scores (RR, 0.96; 95% CI, 0.70–1.31) or more likely to experience moderate or severe anxiety and depression (RR, 1.14; 95% CI, 0.65–2.02). Similar findings were observed when analysis was restricted to women with abnormal cytology or positive hrHPV test results. The likelihood of having abnormal long-term anxiety or depression scores for 4–24 months after screening in women aged 34 years and older was not affected by the screening method or the screening results.

Although anxiety and distress associated with screening and diagnosis have been reported, findings differed in studies because of sociodemographic, behavioural, and age differences in women included in these studies. In a qualitative study in Denmark examining the experiences of women with different stages of cervical dysplasia and whether their knowledge of HPV as the cause of cervical dysplasia influenced their perception of their disease, Lee Mortensen & Adeler (2010) conducted a focus group interview of 12 women with different stages of cervical dysplasia. The participants considered cervical dysplasia to be a highly distressing condition and experienced monitoring before regression of the lesions or treatment could be initiated as a worrying delay. Women expressed a fear of cancer that was not proportional to the stage of their dysplasia, but was determined by their degree of knowledge about their condition. The results suggested that although physicians are the source of information for patients, women's concerns were dependent on the quality of communication with medical practitioners and the amount of information provided.

(v) Follow-up because of an abnormal cytology result

Women with abnormal test results can be monitored by repeat cytological procedures or HPV testing after initial diagnosis (see Section 4.4.8 for HPV testing follow-up).

Kitchener et al. (2004) conducted an RCT of women attending routine screening and with recurrent BMD smear results in the United Kingdom, to determine whether a choice between colposcopy or cytological surveillance at 6 months would be beneficial to women with mildly abnormal smears in terms of psychological morbidity when compared with the national policy of surveillance at 6 months. Women were assigned to either a repeat cytology group ($n = 243$) or a choice group, in which they could choose between repeat cytology and colposcopy ($n = 233$). A survey of psychological effects was then undertaken using the GHQ and STAI questionnaires. Questionnaires were completed at baseline and repeated after initial colposcopy, if chosen, and again before and after the visit at 6 months (cytology or colposcopy) and finally at 12 months. Mean scores for GHQ and STAI state anxiety levels were no different between

the choice and no-choice groups. Both general health scores on GHQ and STAI state anxiety levels decreased over 12 months in both groups, whatever the strategy.

In the TOMBOLA trial, 3399 women aged 20–59 years with low-grade cytological abnormalities detected in the NHS Cervical Screening Programme in the United Kingdom were randomized to cytological surveillance or initial colposcopy and invited to complete a psychological questionnaire survey at recruitment and at 12, 18, 24, and 30 months. Over 30 months, women assigned to the colposcopy arm had lower scores for worries related to follow-up compared with women assigned to the cytology surveillance arm (Fielding et al., 2017). Women assigned to the colposcopy group reported lower levels of satisfaction with information and support than women assigned to the cytology surveillance group.

In a study in 1555 women aged 20–59 years referred for colposcopy after a low-grade cytology result and followed up for 30 months, 40% of women worried about having cervical cancer at one or more time point during follow-up, 26% worried about having sex, 24% worried about future fertility, and 60% worried about their general health (Sharp et al., 2015). Women diagnosed with CIN2+ had significantly higher risks of worries about cervical cancer and future fertility, and the management received was significantly associated with worries about cervical cancer and having sex. Younger women more often reported worries about future fertility.

References

Aimagambetova G, Azizan A (2018). Epidemiology of HPV infection and HPV-related cancers in Kazakhstan: a review. *Asian Pac J Cancer Prev*. 19(5):1175–80. doi:10.22034/APJCP.2018.19.5.1175 PMID:29801397

Akamatsu S, Kodama S, Himeji Y, Ikuta N, Shimagaki N (2012). A comparison of liquid-based cytology with conventional cytology in cervical cancer screening. *Acta Cytol*. 56(4):370–4. doi:10.1159/000337641 PMID:22846595

Aktanko AP, Boytsova TYU, Klimkovich NM (2018). Laboratory cytology screening results of the cervix smear-prints as a marker of intraepithelial pathological lesions. [in Russian] *Health Care of the Far East*. 2(76):51–54. Available from: https://www.elibrary.ru/item.asp?id=35586309.

Albrow R, Kitchener H, Gupta N, Desai M (2012). Cervical screening in England: the past, present, and future. *Cancer Cytopathol*. 120(2):87–96. doi:10.1002/cncy.20203 PMID:22367883

Allia E, Ronco G, Coccia A, Luparia P, Macrì L, Fiorito C, et al. (2015). Interpretation of p16^{INK4a}/Ki-67 dual immunostaining for the triage of human papillomavirus-positive women by experts and nonexperts in cervical cytology. *Cancer Cytopathol*. 123(4):212–8. doi:10.1002/cncy.21511 PMID:25534957

Almonte M, Ferreccio C, Winkler JL, Cuzick J, Tsu V, Robles S, et al. (2007). Cervical screening by visual inspection, HPV testing, liquid-based and conventional cytology in Amazonian Peru. *Int J Cancer*. 121(4):796–802. doi:10.1002/ijc.22757 PMID:17437272

Andrae B, Kemetli L, Sparén P, Silfverdal L, Strander B, Ryd W, et al. (2008). Screening-preventable cervical cancer risks: evidence from a nationwide audit in Sweden. *J Natl Cancer Inst*. 100(9):622–9. doi:10.1093/jnci/djn099 PMID:18445828

Andreassen T, Hansen BT, Engesæter B, Hashim D, Støer NC, Tropé A, et al. (2019). Psychological effect of cervical cancer screening when changing primary screening method from cytology to high-risk human papilloma virus testing. *Int J Cancer*. 145(1):29–39. doi:10.1002/ijc.32067 PMID:30549273

Anttila A, Pokhrel A, Kotaniemi-Talonen L, Hakama M, Malila N, Nieminen P (2011). Cervical cancer patterns with automation-assisted and conventional cytological screening: a randomized study. *Int J Cancer*. 128(5):1204–12. doi:10.1002/ijc.25677 PMID:20848590

Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al., editors (2008a). European guidelines for quality assurance in cervical cancer screening. 2nd ed. Luxembourg: Office for Official Publications of the European Communities.

Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J (2008b). Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. *Obstet Gynecol*. 111(1):167–77. doi:10.1097/01.AOG.0000296488.85807.b3 PMID: 18165406

Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. (2020). Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health*. 8(2):e191–203. doi:10.1016/S2214-109X(19)30482-6 PMID:31812369

Aristizabal N, Cuello C, Correa P, Collazos T, Haenszel W (1984). The impact of vaginal cytology on cervical cancer risks in Cali, Colombia. *Int J Cancer*. 34(1):5–9. doi:10.1002/ijc.2910340103 PMID:6746118

Austin RM, Ramzy I (1998). Increased detection of epithelial cell abnormalities by liquid-based gynecologic cytology preparations. A review of accumulated data. *Acta Cytol*. 42(1):178–84. doi:10.1159/000331543 PMID:9479337

Bain BJ (2017). Preparation and staining methods for blood and bone marrow films. In: Bain BJ, Bates I, Laffan MA, editors. Dacie and Lewis practical haematology. 12th ed. London, UK: Elsevier; pp. 50–60.

Barchuk A, Bespalov A, Huhtala H, Chimed T, Laricheva I, Belyaev A, et al. (2018). Breast and cervical cancer incidence and mortality trends in Russia 1980–2013. *Cancer Epidemiol*. 55:73–80. doi:10.1016/j.canep.2018.05.008 PMID:29843073

Beerman H, van Dorst EBL, Kuenen-Boumeester V, Hogendoorn PCW (2009). Superior performance of liquid-based versus conventional cytology in a population-based cervical cancer screening program. *Gynecol Oncol*. 112(3):572–6. doi:10.1016/j.ygyno.2008.12.012 PMID:19150573

Bekmukhambetov YZ, Mynbaev OA, Tinelli A, Balmagambetova SK, Koyshybaev AK, Urazayev ON, et al. (2018). Human papillomavirus related issues in western Kazakhstan: protocol for a comprehensive study. *Russian Open Medical Journal*. 7(4):e0408. doi:10.15275/rusomj.2018.0408

Benevolo M, Allia E, Gustinucci D, Rollo F, Bulletti S, Cesarini E, et al.; New Technologies for Cervical Cancer Screening 2 (NTCC2) Working Group (2017). Interobserver reproducibility of cytologic $p16^{INK4a}$/Ki-67 dual immunostaining in human papillomavirus-positive women. *Cancer Cytopathol*. 125(3):212–20. doi:10.1002/cncy.21800 PMID:27926800

Bernstein SJ, Sanchez-Ramos L, Ndubisi B (2001). Liquid-based cervical cytologic smear study and conventional Papanicolaou smears: a metaanalysis of prospective studies comparing cytologic diagnosis and sample adequacy. *Am J Obstet Gynecol*. 185(2):308–17. doi:10.1067/mob.2001.116736 PMID:11518884

Berrino F, Gatta G, D'Alto M, Crosignani P, Riboli E (1986). Efficacy of screening in preventing invasive cervical cancer: a case-control study in Milan, Italy. In: Hakama M, Miller AB, Day NE, editors. Screening for cancer of the uterine cervix (IARC Scientific Publications, No. 76). Lyon, France: International Agency for Research on Cancer; pp. 111–23.

Bezrukov AV (2017). Romanowsky staining, the Romanowsky effect and thoughts on the question of scientific priority. *Biotech Histochem*. 92(1):29–35. doi:10.1080/10520295.2016.1250285 PMID:28098484

Biscotti CV, Dawson AE, Dziura B, Galup L, Darragh T, Rahemtulla A, et al. (2005). Assisted primary screening using the automated ThinPrep Imaging System. *Am J Clin Pathol*. 123(2):281–7. doi:10.1309/AGB1MJ9H5N43MEGX PMID:15842055

Blanks RG, Kelly RS (2010). Comparison of cytology and histology results in English cervical screening laboratories before and after liquid-based cytology conversion: do the data provide evidence for a single category of high-grade dyskaryosis? *Cytopathology*. 21(6):368–73. doi:10.1111/j.1365-2303.2010.00808.x PMID:20955376

Bloomfield HE, Olson A, Greer N, Cantor A, MacDonald R, Rutks I, et al. (2014). Screening pelvic examinations in asymptomatic, average-risk adult women: an evidence report for a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 161(1):46–53. doi:10.7326/M13-2881 PMID:24979449

Boon ME, Drijver JS (1986). Routine cytological staining techniques: theoretical background and practice. London, UK: Palgrave. doi:10.1007/978-1-349-18250-3

Boon ME, Tabbers-Bouwmeester ML (1980). Gynaecological cytology: a textbook and atlas. London, UK: Macmillan.

Branca M, Longatto-Filho A (2015). Recommendations on quality control and quality assurance in cervical cytology. *Acta Cytol*. 59(5):361–9. doi:10.1159/000441515 PMID:26569109

Bray F, Loos AH, McCarron P, Weiderpass E, Arbyn M, Møller H, et al. (2005). Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer Epidemiol Biomarkers Prev*. 14(3):677–86. doi:10.1158/1055-9965.EPI-04-0569 PMID:15767349

Broder S (1992). The Bethesda System for reporting cervical/vaginal cytologic diagnoses: report of the 1991 Bethesda Workshop. *JAMA*. 267(14):1892. doi:10.1001/jama.1992.03480140014005 PMID:1548812

Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, et al. (2019a). Human papillomavirus and related diseases in Europe. Summary report. Barcelona, Spain: ICO/IARC Information Centre on HPV and Cancer. Available from: https://hpvcentre.net/statistics/reports/XEX.pdf.

Bruni L, Albero G, Serrano B, Mena M, Gómez D, Muñoz J, et al. (2019b). Human papillomavirus and related diseases in Asia. Summary report. Barcelona, Spain: ICO/IARC Information Centre on HPV and Cancer. Available from: https://hpvcentre.net/statistics/reports/XSX.pdf.

Burghardt E (1970). The histomorphologic criteria of beginning stromal invasion in the squamous cell carcinoma of the cervix. *Minerva Ginecol*. 22(23):1142–7. PMID:4930076

Cancer Council Australia Cervical Cancer Screening Guidelines Working Party (2020). National Cervical Screening Program: guidelines for the management of screen-detected abnormalities, screening in specific populations and investigation of abnormal vaginal bleeding. Available from: https://wiki.cancer.org.au/australia/Guidelines:Cervical_cancer/Screening.

Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, van de Pol A, van Krieken JHJM, et al. (2007). Pathology databanking and biobanking in the Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol*. 29(1):19–24. PMID:17429138

Castañón A, Landy R, Cuzick J, Sasieni P (2014). Cervical screening at age 50–64 years and the risk of cervical cancer at age 65 years and older: population-based case control study. *PLoS Med*. 11(1):e1001585. doi:10.1371/journal.pmed.1001585 PMID:24453946

CDC (1997). Regulatory closure of cervical cytology laboratories: recommendations for a public health response. *MMWR Recomm Rep*. 46(RR-17):1–19. PMID:9409538

Chen C, Yang Z, Li Z, Li L (2012). Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis. *Int J Gynecol Cancer*. 22(6):908–21. doi:10.1097/IGC.0b013e318256e5e4 PMID:22672987

Chernyakova AE (2016). Significance of cervical screening in diagnostics and treatment of cervical pathology. [thesis]. [in Russian] Available from: http://repo.knmu.edu.ua/bitstream/123456789/13031/1/2016_01_%D0%9C%D0%A2%D0%A2_%D0%A7%D0%B5%D1%80%D0%BD%D1%8F%D0%BA%D0%BE%D0%B2%D0%B0.pdf.

Cheung AN, Szeto EF, Leung BSY, Khoo US, Ng AWY (2003). Liquid-based cytology and conventional cervical smears: a comparison study in an Asian screening population. *Cancer*. 99(6):331–5. doi:10.1002/cncr.11786 PMID:14681939

Chichareon SB, Tassee S, Wootipoom V, Buhachat R, Harnprasertpong J (2005). Situation analysis for management of abnormal Pap smears in the lower southern Thailand. *Asian Pac J Cancer Prev*. 6(3):286–94. PMID:16235988

Chichareon SB, Tocharoenvanich ST (2002). Situation analysis of atypical squamous cells of undetermined significance smears in Songklanagarind Hospital. *J Obstet Gynaecol Res*. 28(5):269–73. doi:10.1046/j.1341-8076.2002.00054.x PMID:12428697

Clarke EA, Anderson TW (1979). Does screening by "Pap" smears help prevent cervical cancer? A case-control study. *Lancet*. 314(8132):1–4. https://di.org/10.1016/S0140-6736(79)90172-7doi:10.1016/S0140-6736(79)90172-7 PMID:87887

Colgan TJ, McLachlin CM, Cotterchio M, Howlett R, Seidenfeld AM, Mai VM (2004). Results of the implementation of liquid-based cytology – SurePath in the Ontario screening program. *Cancer*. 102(6):362–7. doi:10.1002/cncr.20656 PMID:15481083

College of American Pathologists (2014). Cytopathology checklist. Northfield (IL), USA: College of American Pathologists. Available from: https://webapps.cap.org/apps/docs/education/OnlineCourseContent/2014/TLTM/CYP04212014.PDF.

Colletti SM, Tranesh GA, Nassar A (2017). Significance of finding benign endometrial cells in women 40–45 versus 46 years or older on Papanicolaou tests and histologic follow-up. *Cytojournal*. 14:22. doi:10.4103/cytojournal.cytojournal_16_17 PMID:29021811

Confortini M, Bondi A, Cariaggi MP, Carozzi F, Dalla Palma P, Ghiringhello B, et al. (2007). Interlaboratory reproducibility of liquid-based equivocal cervical cytology within a randomized controlled trial framework. *Diagn Cytopathol*. 35(9):541–4. doi:10.1002/dc.20681 PMID:17703457

Confortini M, Bulgaresi P, Cariaggi MP, Carozzi FM, Cecchini S, Cipparrone I, et al. (2004). Comparing conventional and liquid-based smears from a consecutive series of 297 subjects referred to colposcopy assessment. *Cytopathology*. 15(3):168–70. doi:10.1111/j.1365-2303.2004.00142.x PMID:15165276

Cooper CP, Saraiya M (2014). Opting out of cervical cancer screening: physicians who do not perform Pap tests. *Am J Prev Med*. 47(3):315–9. doi:10.1016/j.amepre.2014.04.018 PMID:24953518

Coste J, Cochand-Priollet B, de Cremoux P, Le Galès C, Cartier I, Molinié V, et al.; French Society of Clinical Cytology Study Group (2003). Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. *BMJ*. 326(7392):733. doi:10.1136/bmj.326.7392.733 PMID:12676841

Cox JT (2009). History of the use of HPV testing in cervical screening and in the management of abnormal cervical screening results. *J Clin Virol*. 45(Suppl 1):S3–12. doi:10.1016/S1386-6532(09)70002-2 PMID:19651367

Cuschieri K, Ronco G, Lorincz A, Smith L, Ogilvie G, Mirabello L, et al. (2018). Eurogin roadmap 2017: triage strategies for the management of HPV-positive women in cervical screening programs. *Int J Cancer*. 143(4):735–45. doi:10.1002/ijc.31261 PMID:29341110

Davey E, Barratt A, Irwig L, Chan SF, Macaskill P, Mannes P, et al. (2006). Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: a systematic review. *Lancet*. 367(9505):122–32. doi:10.1016/S0140-6736(06)67961-0 PMID:16413876

Davey E, d'Assuncao J, Irwig L, Macaskill P, Chan SF, Richards A, et al. (2007). Accuracy of reading liquid based cytology slides using the ThinPrep Imager compared with conventional cytology: prospective study. *BMJ*. 335(7609):31. doi:10.1136/bmj.39219.645475.55 PMID:17604301

Davies P, Gentili C, Valuta D, Petrovici V, Caras O, Cuznețova T, et al. (2016). Assessment of cervical cytology services in the Republic of Moldova and recommendations for their reorganisation to comply with international evidence-based standards. Chisinau, Moldova: Ministry of Health of Moldova and United Nations Population Fund. Available from: https://moldova.unfpa.org/sites/default/files/pub-pdf/RM%20Cervical%20Cytology%20Assessment%20%20Recommendations_eng.pdf.

de Morais LSF, Magalhães JC, Braga IDS, Marega LA, Tavares SBDN, Amaral RG (2020). Performance of laboratories after 10 years of participating in external quality monitoring in cervical cytology. *Acta Cytol*. 64(3):224–31. doi:10.1159/000502433 PMID:31480038

Decker K, Demers A, Chateau D, Musto G, Nugent Z, Lotocki R, et al. (2009). Papanicolaou test utilization and frequency of screening opportunities among women diagnosed with cervical cancer. *Open Med*. 3(3):e140–7. PMID:21603052

Dickinson JA, Stankiewicz A, Popadiuk C, Pogany L, Onysko J, Miller AB (2012). Reduced cervical cancer incidence and mortality in Canada: national data from 1932 to 2006. *BMC Public Health*. 12(1):992. doi:10.1186/1471-2458-12-992 PMID:23158654

Doornewaard H, van der Schouw YT, van der Graaf Y, Bos AB, Habbema JD, van den Tweel JG (1999). The diagnostic value of computer-assisted primary cervical smear screening: a longitudinal cohort study. *Mod Pathol*. 12(11):995–1000. PMID:10574595

Duggan MA (2000). Papnet-assisted, primary screening of cervico-vaginal smears. *Eur J Gynaecol Oncol*. 21(1):35–42. PMID:10726616

Dugué PA, Lynge E, Rebolj M (2014). Mortality of non-participants in cervical screening: register-based cohort study. *Int J Cancer*. 134(11):2674–82. doi:10.1002/ijc.28586 PMID:24288241

Elit L (2014). Role of cervical screening in older women. *Maturitas*. 79(4):413–20. doi:10.1016/j.maturitas.2014.09.012 PMID:25449664

Elsheikh TM, Austin RM, Chhieng DF, Miller FS, Moriarty AT, Renshaw AA (2013). American Society of Cytopathology workload recommendations for automated Pap test screening: developed by the Productivity and Quality Assurance in the Era of Automated Screening Task Force. *Diagn Cytopathol*. 41(2):174–8. doi:10.1002/dc.22817 PMID:22351120

Farnsworth A (2016). Cervical cancer screening in Australia: past and present. *Cancer Cytopathol*. 124(4):231–4. doi:10.1002/cncy.21642 PMID:26513423

Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, et al. (2018). Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer*. 103:356–87. doi:10.1016/j.ejca.2018.07.005 PMID:30100160

Fiddes P, Scott A, Fletcher J, Glasier A (2003). Attitudes towards pelvic examination and chaperones: a questionnaire survey of patients and providers. *Contraception*. 67(4):313–7. doi:10.1016/S0010-7824(02)00540-1 PMID:12684154

Fielding S, Rothnie K, Gray NM, Little J, Cruickshank ME, Neal K, et al.; TOMBOLA group (2017). Psychosocial morbidity in women with abnormal cervical cytology managed by cytological surveillance or initial colposcopy: longitudinal analysis from the TOMBOLA randomised trial. *Psychooncology*. 26(4):476–83. doi:10.1002/pon.4163 PMID:27297097

Fokom-Domgue J, Combescure C, Fokom-Defo V, Tebeu PM, Vassilakos P, Kengne AP, et al. (2015). Performance of alternative strategies for primary cervical cancer screening in sub-Saharan Africa: systematic review and meta-analysis of diagnostic test accuracy studies. *BMJ*. 351:h3084. doi:10.1136/bmj.h3084 PMID:26142020

Freijomil-Vázquez C, Gastaldo D, Coronado C, Movilla-Fernández MJ (2019). Health care informational challenges for women diagnosed with cervical intraepithelial neoplasia: a qualitative study. *BMC Womens Health*. 19(1):112. doi:10.1186/s12905-019-0811-5 PMID:31477083

Fremont-Smith M, Marino J, Griffin B, Spencer L, Bolick D (2004). Comparison of the SurePath™ liquid-based Papanicolaou smear with the conventional Papanicolaou smear in a multisite direct-to-vial study. *Cancer*. 102(5):269–79. doi:10.1002/cncr.20599 PMID:15386329

French DP, Maissi E, Marteau TM (2004). Psychological costs of inadequate cervical smear test results. *Br J Cancer*. 91(11):1887–92. doi:10.1038/sj.bjc.6602224 PMID:15534608

Gage JC, Ferreccio C, Gonzales M, Arroyo R, Huivín M, Robles SC (2003). Follow-up care of women with an abnormal cytology in a low-resource setting. *Cancer Detect Prev*. 27(6):466–71. doi:10.1016/j.cdp.2003.09.004 PMID:14642555

Garau M, Musetti C, Alonso R, Barrios E (2019). Trends in cancer incidence in Uruguay: 2002–2015. *Colomb Med (Cali)*. 50(4):224–38. doi:10.25100/cm.v50i4.4212 PMID:32476689

Giorgi-Rossi P, Franceschi S, Ronco G (2012). HPV prevalence and accuracy of HPV testing to detect high-grade cervical intraepithelial neoplasia. *Int J Cancer*. 130(6):1387–94. doi:10.1002/ijc.26147 PMID:21520039

Grada Z, Paquette C, Eklund CM, Zhang C, Sung CJ, Steinhoff M, et al. (2017). Evaluating the age cutoff criterion for reporting benign-appearing endometrial cells in routine Pap tests: an 8-year retrospective review. *Acta Cytol*. 61(3):194–8. doi:10.1159/000473694 PMID:28486237

Gray NM, Sharp L, Cotton SC, Masson LF, Little J, Walker LG, et al.; TOMBOLA group (2006). Psychological effects of a low-grade abnormal cervical smear test result: anxiety and associated factors. *Br J Cancer*. 94(9):1253–62. doi:10.1038/sj.bjc.6603086 PMID:16622462

Grebenkina EV, Guryanova AV, Ushakova MS, Belova SV, Smetanina SV, Zhelezin OV (2018). The work of the Reference Cytological Center, including the early diagnosis of cervical cancer in the Nizhny Novgorod Region. [in Russian] *Oncol Bull Volga Region*. 9(3).

Grjibovski AM, Dubovichenko D, Saduakassova S, Zhatkanbayeva G, Omarova G, Shalgumbayeva G, et al. (2018). Incidence, mortality and determinants of survival from cervical cancer in northwest Russia: a registry-based cohort study. *Int Health*. 10(2):92–9. doi:10.1093/inthealth/ihx068 PMID:29432549

Gupta R, Gupta S, Mehrotra R, Sodhani P (2017). Cervical cancer screening in resource-constrained countries: current status and future directions. *Asian Pac J Cancer Prev*. 18(6):1461–7. doi:10.22034/apjcp.2017.18.6.1461 PMID:28669152

Gustinucci D, Giorgi Rossi P, Cesarini E, Broccolini M, Bulletti S, Carlani A, et al. (2016). Use of cytology, E6/E7 mRNA, and $p16^{INK4a}$-Ki-67 to define the management of human papillomavirus (HPV)-positive women in cervical cancer screening. *Am J Clin Pathol*. 145(1):35–45. doi:10.1093/ajcp/aqv019 PMID:26712869

Halford JA, Batty T, Boost T, Duhig J, Hall J, Lee C, et al. (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(5):318–26. doi:10.1002/dc.21199 PMID:19813268

Herbert A, Anic V, Cochand-Priollet B, Dina R, Ehya H, Eide ML, et al. (2014). Training and practice of cytotechnologists: a discussion forum focused on Europe. *Cytopathology*. 25(5):307–15. doi:10.1111/cyt.12201 PMID:25209399

Herbert A, Stein K, Bryant TN, Breen C, Old P (1996). Relation between the incidence of invasive cervical cancer and the screening interval: is a five year interval too long? *J Med Screen*. 3(3):140–5. doi:10.1177/096914139600300307 PMID:8946309

Hernández-Avila M, Lazcano-Ponce EC, de Ruíz PA, Romieu I (1998). Evaluation of the cervical cancer screening programme in Mexico: a population-based case-control study. *Int J Epidemiol*. 27(3):370–6. doi:10.1093/ije/27.3.370 PMID:9698122

Herrero R, Brinton LA, Reeves WC, Brenes MM, de Britton RC, Gaitan E, et al. (1992). Screening for cervical cancer in Latin America: a case-control study. *Int J Epidemiol*. 21(6):1050–6. doi:10.1093/ije/21.6.1050 PMID:1336485

Herrero R, Brinton LA, Reeves WC, Brenes MM, Tenorio F, de Britton RC, et al. (1990). Risk factors for invasive carcinoma of the uterine cervix in Latin America. *Bull Pan Am Health Organ*. 24(3):263–83. PMID:2171706

Hilgarth M (2001). Munich II nomenclature for gynaecological cytodiagnosis of the cervix. *Frauenarzt*. 39:392–3.

Hinson S, Molberg K, Mir M, Flores M, Zheng W, Lucas E (2019). Age cutoff for reporting of benign-appearing endometrial cells in Papanicolaou specimens; should it be raised? A 10-year retrospective study from a large county hospital. *J Am Soc Cytopathol*. 8(2):78–83. doi:10.1016/j.jasc.2018.09.002 PMID:31287423

Hoffman M, Cooper D, Carrara H, Rosenberg L, Kelly J, Stander I, et al. (2003). Limited Pap screening associated with reduced risk of cervical cancer in South Africa. *Int J Epidemiol*. 32(4):573–7. doi:10.1093/ije/dyg081 PMID:12913031

Horobin RW (2011). How Romanowsky stains work and why they remain valuable – including a proposed universal Romanowsky staining mechanism and a rational troubleshooting scheme. *Biotech Histochem*. 86(1):36–51. doi:10.3109/10520295.2010.515491 PMID:21235292

Hosono S, Terasawa T, Katayama T, Sasaki S, Hoshi K, Hamashima C (2018). Frequency of unsatisfactory cervical cytology smears in cancer screening of Japanese women: a systematic review and meta-analysis. *Cancer Sci*. 109(4):934–43. doi:10.1111/cas.13549 PMID:29478281

Hoyo C, Yarnall KSH, Skinner CS, Moorman PG, Sellers D, Reid L (2005). Pain predicts non-adherence to Pap smear screening among middle-aged African American women. *Prev Med*. 41(2):439–45. doi:10.1016/j.ypmed.2004.11.021 PMID:15917039

Hu Y, Hong Z, Gu L, Xie L, Yang B, Dai H, et al. (2020). Evaluation of p16/Ki-67 dual-stained cytology in triaging HPV-positive women during cervical cancer screening. *Cancer Epidemiol Biomarkers Prev*. 29(6):1246–52. doi:10.1158/1055-9965.EPI-19-1180 PMID:32156721

IARC (1986). Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies. IARC Working Group on evaluation of cervical cancer screening programmes. *Br Med J (Clin Res*

Ed). 293(6548):659–64. doi:10.1136/bmj.293.6548.659 PMID:3092971

IARC (2005). Cervix cancer screening. *IARC Handb Cancer Prev*. 10:1–302. Available from: https://publications.iarc.fr/380.

IARC (2012). Current status and future directions of breast and cervical cancer prevention and early detection in Belarus (IARC Working Group Reports, Volume 6). Lyon, France: International Agency for Research on Cancer. Available from: https://publications.iarc.fr/389.

Ideström M, Milsom I, Andersson-Ellström A (2003). Women's experience of coping with a positive Pap smear: a register-based study of women with two consecutive Pap smears reported as CIN 1. *Acta Obstet Gynecol Scand*. 82(8):756–61.doi:10.1034/j.1600-0412.2003.00165.x PMID:12848648

Ideström M, Milsom I, Andersson-Ellström A (2007). The cervical cancer screening program from a midwife's perspective. *Acta Obstet Gynecol Scand*. 86(6):742–8. doi:10.1080/00016340701371272 PMID:17520410

Ikenberg H, Bergeron C, Schmidt D, Griesser H, Alameda F, Angeloni C, et al.; PALMS Study Group (2013). Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: results of the PALMS study. *J Natl Cancer Inst*. 105(20):1550–7. doi:10.1093/jnci/djt235 PMID:24096620

International Collaboration of Epidemiological Studies of Cervical Cancer (2007). Comparison of risk factors for invasive squamous cell carcinoma and adenocarcinoma of the cervix: collaborative reanalysis of individual data on 8,097 women with squamous cell carcinoma and 1,374 women with adenocarcinoma from 12 epidemiological studies. *Int J Cancer*. 120(4):885–91. doi:10.1002/ijc.22357 PMID:17131323

Iskhakova GS, Imelbaeva EA, Amineva RKh (2012). Experience of a centralized cytological laboratory in the central regional hospital of the small town Meleuz. [in Russian] *Clinical Cytology News in Russia*. 16(1–2).

Ito K, Kimura R, Konishi H, Ozawa N, Yaegashi N, Ohashi Y, et al. (2020). A comparison of liquid-based and conventional cytology using data for cervical cancer screening from the Japan Cancer Society. *Jpn J Clin Oncol*. 50(2):138–44. doi:10.1093/jjco/hyz161 PMID:31735963

Jarynowski A (2019). Cost-effectiveness analysis for HPV mitigation strategies in the Republic of Moldova based on infectious disease modelling. *Econ Sociol*. 2:50–64. doi:10.36004/nier.es.2019.2-05

Jiang MY, Wu Z, Li T, Yu L, Zhang SK, Zhang X, et al. (2020). Performance of HPV genotyping combined with p16/Ki-67 in detection of cervical precancer and cancer among HPV-positive Chinese women. *Cancer Prev Res (Phila)*. 13(2):163–72. doi:10.1158/1940-6207.CAPR-19-0144 PMID:31871224

Jiménez-Pérez M, Thomas DB (1999). Has the use of Pap smears reduced the risk of invasive cervical cancer in Guadalajara, Mexico? *Int J Cancer*. 82(6):804–9. doi:10.1002/(SICI)1097-0215(19990909)82:6<804::AID-IJC6>3.0.CO;2-N PMID:10446445

Jun JK, Choi KS, Jung KW, Lee HY, Gapstur SM, Park EC, et al. (2009). Effectiveness of an organized cervical cancer screening program in Korea: results from a cohort study. *Int J Cancer*. 124(1):188–93. doi:10.1002/ijc.23841 PMID:18785204

Kahn JA, Goodman E, Huang B, Slap GB, Emans SJ (2003). Predictors of Papanicolaou smear return in a hospital-based adolescent and young adult clinic. *Obstet Gynecol*. 101(3):490–9. doi:10.1016/S0029-7844(02)02592-9 PMID:12636952

Kamineni A, Weinmann S, Shy KK, Glass AG, Weiss NS (2013). Efficacy of screening in preventing cervical cancer among older women. *Cancer Causes Control*. 24(9):1653–60. doi:10.1007/s10552-013-0239-4 PMID:23744043

Kasinpila C, Promthet S, Vatanasapt P, Sasieni P, Parkin DM (2011). Evaluation of the nationwide cervical screening programme in Thailand: a case-control study. *J Med Screen*. 18(3):147–53. doi:10.1258/jms.2011.011075 PMID:22045824

Kirillina MP, Ivanova AK, Lushnikova EL (2018). Cytological analysis of cervix inflammatory diseases in different age groups on the example of the Sakha (Yakutia) Republic. *Yakut Medical Journal*. 4(64):12–5. doi:10.25789/YMJ.2018.64.03

Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, et al. (2009). HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol*. 10(7):672–82. doi:10.1016/S1470-2045(09)70156-1 PMID:19540162

Kitchener HC, Blanks R, Dunn G, Gunn L, Desai M, Albrow R, et al. (2011). Automation-assisted versus manual reading of cervical cytology (MAVARIC): a randomised controlled trial. *Lancet Oncol*. 12(1):56–64. doi:10.1016/S1470-2045(10)70264-3 PMID:21146458

Kitchener HC, Burns S, Nelson L, Myers AJ, Fletcher I, Desai M, et al. (2004). A randomised controlled trial of cytological surveillance versus patient choice between surveillance and colposcopy in managing mildly abnormal cervical smears. *BJOG*. 111(1):63–70. doi:10.1046/j.1471-0528.2003.00007.x PMID:14687054

Kituncharoen S, Tantbirojn P, Niruthisard S (2015). Comparison of unsatisfactory rates and detection of abnormal cervical cytology between conventional Papanicolaou smear and liquid-based cytology (Sure Path®). *Asian Pac J Cancer Prev*. 16(18):8491–4. doi:10.7314/apjcp.2015.16.18.8491 PMID:26745107

Klaes R, Friedrich T, Spitkovsky D, Ridder R, Rudy W, Petry U, et al. (2001). Overexpression of $p16^{INK4A}$ as a specific marker for dysplastic and neoplastic epithelial

cells of the cervix uteri. *Int J Cancer*. 92(2):276–84. doi:10.1002/ijc.1174 PMID:11291057

Klug SJ, Neis KJ, Harlfinger W, Malter A, König J, Spieth S, et al. (2013). A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. *Int J Cancer*. 132(12):2849–57. doi:10.1002/ijc.27955 PMID:23175270

Kok MR, Boon ME (1996). Consequences of neural network technology for cervical screening: increase in diagnostic consistency and positive scores. *Cancer*. 78(1):112–7. doi:10.1002/(SICI)1097-0142(19960701)78:1<112::AID-CNCR16>3.0.CO;2-2 PMID:8646706

Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PPL, Mustafa RA, et al. (2017). Cytology versus HPV testing for cervical cancer screening in the general population. *Cochrane Database Syst Rev*. 8:CD008587. doi:10.1002/14651858.CD008587.pub2 PMID:28796882

Korfage IJ, van Ballegooijen M, Huveneers H, Essink-Bot ML (2010). Anxiety and borderline Pap smear results. *Eur J Cancer*. 46(1):134–41. doi:10.1016/j.ejca.2009.07.003 PMID:19683428

Korfage IJ, van Ballegooijen M, Wauben B, Looman CWN, Habbema JDF, Essink-Bot ML (2012). Having a Pap smear, quality of life before and after cervical screening: a questionnaire study. *BJOG*. 119(8):936–44. doi:10.1111/j.1471-0528.2012.03344.x PMID:22568482

Kozyreva SM, et al. (2012). On the clinical and morphological diagnosis of cervical neoplasms by ploidometry. [in Russian] *Bull New Med Technol*. 14(1):137.

Krafts KP, Pambuccian SE (2011). Romanowsky staining in cytopathology: history, advantages and limitations. *Biotech Histochem*. 86(2):82–93. doi:10.3109/10520295.2010.515492 PMID:21395493

Kurtycz DFI, Staats PN, Chute DJ, Russell D, Pavelec D, Monaco SE, et al. (2017). Bethesda Interobserver Reproducibility Study-2 (BIRST-2): Bethesda System 2014. *J Am Soc Cytopathol*. 6(4):131–44. doi:10.1016/j.jasc.2017.03.003 PMID:31043266

La Vecchia C, Bosetti C, Lucchini F, Bertuccio P, Negri E, Boyle P, et al. (2010). Cancer mortality in Europe, 2000–2004, and an overview of trends since 1975. *Ann Oncol*. 21(6):1323–60. doi:10.1093/annonc/mdp530 PMID:19948741

La Vecchia C, Franceschi S, Decarli A, Fasoli M, Gentile A, Tognoni G (1984). "Pap" smear and the risk of cervical neoplasia: quantitative estimates from a case-control study. *Lancet*. 342(8406):779–82. doi:10.1016/S0140-6736(84)90705-0 PMID:6148523

Larsen M, Oldeide CC, Malterud K (1997). Not so bad after all... Women's experiences of pelvic examinations. *Fam Pract*. 14(2):148–52. doi:10.1093/fampra/14.2.148 PMID:9137954

Lee Mortensen G, Adeler AL (2010). Qualitative study of women's anxiety and information needs after a diagnosis of cervical dysplasia. *J Public Health*. 18(5):473–82. doi:10.1007/s10389-010-0330-1 PMID:21151479

Lei J, Andrae B, Ploner A, Lagheden C, Eklund C, Nordqvist Kleppe S, et al. (2019). Cervical screening and risk of adenosquamous and rare histological types of invasive cervical carcinoma: population based nested case-control study. *BMJ*. 365:l1207. doi:10.1136/bmj.l1207 PMID:30944091

Li Q, Weina PJ, Miller RS (2012). Malaria analysis. In: Kottke-Marchant K, Davis BH, editors. Laboratory hematology practice. 1st ed. Oxford, UK: Wiley-Blackwell; pp. 626–37. doi:10.1002/9781444398595.ch48

Lönnberg S, Anttila A, Luostarinen T, Nieminen P (2012). Age-specific effectiveness of the Finnish cervical cancer screening programme. *Cancer Epidemiol Biomarkers Prev*. 21(8):1354–61. doi:10.1158/1055-9965.EPI-12-0162 PMID:22665576

Lönnberg S, Nieminen P, Luostarinen T, Anttila A (2013). Mortality audit of the Finnish cervical cancer screening program. *Int J Cancer*. 132(9):2134–40. doi:10.1002/ijc.27844 PMID:22987437

Maccallini V, Angeloni C, Caraceni D, Fortunato C, Venditti MA, Di Gabriele G, et al. (2008). Comparison of the conventional cervical smear and liquid-based cytology: results of a controlled, prospective study in the Abruzzo region of Italy. *Acta Cytol*. 52(5):568–74. doi:10.1159/000325599 PMID:18833819

Maissi E, Marteau TM, Hankins M, Moss S, Legood R, Gray A (2004). Psychological impact of human papillomavirus testing in women with borderline or mildly dyskaryotic cervical smear test results: cross sectional questionnaire study. *BMJ*. 328(7451):1293. doi:10.1136/bmj.328.7451.1293 PMID:15166066

Makino H, Sato S, Yajima A, Komatsu S, Fukao A (1995). Evaluation of the effectiveness of cervical cancer screening: a case-control study in Miyagi, Japan. *Tohoku J Exp Med*. 175(3):171–8. doi:10.1620/tjem.175.171 PMID:7792786

Maksem JA, Finnemore M, Belsheim BL, Roose EB, Makkapati SR, Eatwell L, et al. (2001). Manual method for liquid-based cytology: a demonstration using 1,000 gynecological cytologies collected directly to vial and prepared by a smear-slide technique. *Diagn Cytopathol*. 25(5):334–8. doi:10.1002/dc.2166 PMID:11747227

Marshall PN (1978). Romanowsky-type stains in haematology. *Histochem J*. 10(1):1–29. doi:10.1007/BF01003411 PMID:74370

Martin-Hirsch PPL, Jarvis GG, Kitchener HC, Lilford R (2000). Collection devices for obtaining cervical cytology samples. *Cochrane Database Syst Rev*. (2):CD001036. doi:10.1002/14651858.CD001036 PMID:10796736

Massad LS, Einstein MH, Huh WK, Katki HA, Kinney WK, Schiffman M, et al.; 2012 ASCCP Consensus Guidelines Conference (2013). 2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors. *Obstet Gynecol*. 121(4):829–46. doi:10.1097/AOG.0b013e3182883a34 PMID:23635684

McCredie MRE, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. (2008). Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol*. 9(5):425–34. doi:10.1016/S1470-2045(08)70103-7 PMID:18407790

McDonald P, Herity B, Johnson Z, O'Kelly F (2001). Views of Irish general practitioners on screening for cervical cancer. *Ir J Med Sci*. 170(3):186–8. doi:10.1007/BF03173887 PMID:12120972

Mendes D, Mesher D, Pista A, Baguelin M, Jit M (2018). Understanding differences in cervical cancer incidence in Western Europe: comparing Portugal and England. *Eur J Public Health*. 28(2):343–7. doi:10.1093/eurpub/ckx176 PMID:29059348

Miller MG, Sung HY, Sawaya GF, Kearney KA, Kinney W, Hiatt RA (2003). Screening interval and risk of invasive squamous cell cervical cancer. *Obstet Gynecol*. 101(1):29–37. doi:10.1016/s0029-7844(02)02454-7 PMID:12517642

Ministry of Health of the Russian Federation (2019). Order of the Ministry of Health of the Russian Federation of March 13, 2019, No 124n. On approval of the procedure for conducting preventive medical examination and clinical examination of certain groups of the adult population. Moscow, Russian Federation: Ministry of Health of the Russian Federation.

Mitchell HS (2001). Longitudinal analysis of histologic high-grade disease after negative cervical cytology according to endocervical status. *Cancer*. 93(4):237–40. doi:10.1002/cncr.9035 PMID:11507695

Mody DR, Davey DD, Branca M, Raab SS, Schenck UG, Stanley MW, et al. (2000). Quality assurance and risk reduction guidelines. *Acta Cytol*. 44(4):496–507. doi:10.1159/000328521 PMID:10934940

Moettus A, Sklar D, Tandberg D (1999). The effect of physician gender on women's perceived pain and embarrassment during pelvic examination. *Am J Emerg Med*. 17(7):635–7. doi:10.1016/S0735-6757(99)90148-1 PMID:10597078

Morrison AS (1992). Screening in chronic disease. Monographs in Epidemiology and Biostatistics. Oxford, UK: Oxford University Press.

Murillo R, Cendales R, Wiesner C, Piñeros M, Tovar S (2009). Effectiveness of cytology-based cervical cancer screening in the Colombian health system. [in Spanish] *Biomedica*. 29(3):354–61. doi:10.7705/biomedica.v29i3.7 PMID:20436987

Mustafa RA, Santesso N, Khatib R, Mustafa AA, Wiercioch W, Kehar R, et al. (2016). Systematic reviews and meta-analyses of the accuracy of HPV tests, visual inspection with acetic acid, cytology, and colposcopy. *Int J Gynaecol Obstet*. 132(3):259–65. doi:10.1016/j.ijgo.2015.07.024 PMID:26851054

Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, et al. (2000). Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med*. 132(10):810–9. doi:10.7326/0003-4819-132-10-200005160-00009 PMID:10819705

Nandini NM, Nandish SM, Pallavi P, Akshatha SK, Chandrashekhar AP, Anjali S, et al. (2012). Manual liquid based cytology in primary screening for cervical cancer – a cost effective preposition for scarce resource settings. *Asian Pac J Cancer Prev*. 13(8):3645–51. doi:10.7314/APJCP.2012.13.8.3645 PMID:23098448

Nascimento MI, Silva GA, Monteiro GT (2012). [Previous history of Pap smears and cervical cancer: a case-control study in the Baixada Fluminense, Rio de Janeiro State, Brazil]. *Cad Saude Publica*. 28(10):1841–53. [Portuguese]doi:10.1590/S0102-311X2012001000004 PMID:23090165

National Institute for Health and Clinical Excellence (2003). Guidance on the use of liquid-based cytology for cervical screening: technology appraisal guidance. London, UK: National Institute for Health and Clinical Excellence. Available from: https://www.nice.org.uk/guidance/TA69.

National Pathology Accreditation Advisory Council (2019). Requirements for laboratories reporting tests for the National Cervical Screening Program (second edition 2019). Canberra, Australia: Australia Department of Health. Available from: https://www1.health.gov.au/internet/main/publishing.nsf/Content/npaac-cervical-screening.

Nayar R, Wilbur DC (2015). The Pap test and Bethesda 2014. *Cancer Cytopathol*. 123(5):271–81. doi:10.1002/cncy.21521 PMID:25931431

Nieminen P, Hakama M, Viikki M, Tarkkanen J, Anttila A (2003). Prospective and randomised public-health trial on neural network-assisted screening for cervical cancer in Finland: results of the first year. *Int J Cancer*. 103(3):422–6. doi:10.1002/ijc.10839 PMID:12471627

Nieminen P, Kallio M, Anttila A, Hakama M (1999). Organised vs. spontaneous Pap-smear screening for cervical cancer: a case-control study. *Int J Cancer*. 83(1):55–8. doi:10.1002/(SICI)1097-0215(19990924)83:1<55::AID-IJC11>3.0.CO;2-U PMID:10449608

Nieminen P, Kotaniemi-Talonen L, Hakama M, Tarkkanen J, Martikainen J, Toivonen T, et al. (2007). Randomized evaluation trial on automation-assisted screening for cervical cancer: results after 777,000 invitations. *J Med Screen*. 14(1):23–8. doi:10.1258/096914107780154468 PMID:17362568

Norrell LL, Kuppermann M, Moghadassi MN, Sawaya GF (2017). Women's beliefs about the purpose and value of routine pelvic examinations. *Am J Obstet Gynecol*. 217(1):86.e1–6. doi:10.1016/j.ajog.2016.12.031 PMID:28040449

Obwegeser JH, Brack S (2001). Does liquid-based technology really improve detection of cervical neoplasia? A prospective, randomized trial comparing the ThinPrep Pap Test with the conventional Pap Test, including follow-up of HSIL cases. *Acta Cytol*. 45(5):709–14. doi:10.1159/000328292 PMID:11575648

Odongua N, Chae YM, Kim MR, Yun JE, Jee SH (2007). Associations between smoking, screening, and death caused by cervical cancer in Korean women. *Yonsei Med J*. 48(2):192–200. doi:10.3349/ymj.2007.48.2.192 PMID:17461516

Ogilvie GS, Krajden M, van Niekerk D, Smith LW, Cook D, Ceballos K, et al. (2017). HPV For Cervical Cancer Screening (HPV FOCAL): complete round 1 results of a randomized trial comparing HPV-based primary screening to liquid-based cytology for cervical cancer. *Int J Cancer*. 140(2):440–8. doi:10.1002/ijc.30454 PMID:27685757

Olson B, Gribble B, Dias J, Curryer C, Vo K, Kowal P, et al. (2016). Cervical cancer screening programs and guidelines in low- and middle-income countries. *Int J Gynaecol Obstet*. 134(3):239–46. doi:10.1016/j.ijgo.2016.03.011 PMID:27350227

Padma S, Ramana PV Kante R, Shravan Kumar O (2018). A comparative study of staining characteristics of Leishman-Geimsa cocktail and Papanicolaou stain in cervical cytology. *Asian Pac. J. Health Sci.* 5(3):233–6. doi:10.21276/apjhs.2018.5.3.32

Pankaj S, Nazneen S, Kumari S, Kumari A, Kumari A, Kumari J, et al. (2018). Comparison of conventional Pap smear and liquid-based cytology: a study of cervical cancer screening at a tertiary care center in Bihar. *Indian J Cancer*. 55(1):80–3. doi:10.4103/ijc.IJC_352_17 PMID:30147099

Pankakoski M, Anttila A, Sarkeala T, Heinävaara S (2019). Effectiveness of cervical cancer screening at age 65 – a register-based cohort study. *PLoS One*. 14(3):e0214486. doi:10.1371/journal.pone.0214486 PMID:30913262

Park Y, Vongdala C, Kim J, Ki M (2015). Changing trends in the incidence (1999–2011) and mortality (1983–2013) of cervical cancer in the Republic of Korea. *Epidemiol Health*. 37:e2015024. doi:10.4178/epih/e2015024 PMID:26063351

Payne N, Chilcott J, McGoogan E (2000). Liquid-based cytology in cervical screening: a rapid and systematic review. *Health Technol Assess*. 4(18):1–73. doi:10.3310/hta4180 PMID:10932023

Pedersen K, Fogelberg S, Thamsborg LH, Clements M, Nygård M, Kristiansen IS, et al. (2018). An overview of cervical cancer epidemiology and prevention in Scandinavia. *Acta Obstet Gynecol Scand*. 97(7):795–807. doi:10.1111/aogs.13313 PMID:29388202

Peeters E, Wentzensen N, Bergeron C, Arbyn M (2019). Meta-analysis of the accuracy of p16 or p16/Ki-67 immunocytochemistry versus HPV testing for the detection of CIN2+/CIN3+ in triage of women with minor abnormal cytology. *Cancer Cytopathol*. 127(3):169–80. doi:10.1002/cncy.22103 PMID:30811902

Peirson L, Fitzpatrick-Lewis D, Ciliska D, Warren R (2013). Screening for cervical cancer: a systematic review and meta-analysis. *Syst Rev*. 2(1):35. doi:10.1186/2046-4053-2-35 PMID:23706117

Petry KU, Menton S, Menton M, van Loenen-Frosch F, de Carvalho Gomes H, Holz B, et al. (2003). Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *Br J Cancer*. 88(10):1570–7. doi:10.1038/sj.bjc.6600918 PMID:12771924

Petry KU, Schmidt D, Scherbring S, Luyten A, Reinecke-Lüthge A, Bergeron C, et al. (2011). Triaging Pap cytology negative, HPV positive cervical cancer screening results with p16/Ki-67 dual-stained cytology. *Gynecol Oncol*. 121(3):505–9. doi:10.1016/j.ygyno.2011.02.033 PMID:21420158

Phaliwong P, Pariyawateekul P, Khuakoonratt N, Sirichai W, Bhamarapravatana K, Suwannarurk K (2018). Cervical cancer detection between conventional and liquid based cervical cytology: a 6-year experience in northern Bangkok Thailand. *Asian Pac J Cancer Prev*. 19(5):1331–6. doi:10.22034/apjcp.2018.19.5.1331 PMID:29802695

Pilleron S, Cabasag CJ, Ferlay J, Bray F, Luciani S, Almonte M, et al. (2020). Cervical cancer burden in Latin America and the Caribbean: where are we? *Int J Cancer*. 147(6):1638–48. doi:10.1002/ijc.32956 PMID:32150288

Pirotta M, Ung L, Stein A, Conway EL, Mast TC, Fairley CK, et al. (2009). The psychosocial burden of human papillomavirus related disease and screening interventions. *Sex Transm Infect*. 85(7):508–13. doi:10.1136/sti.2009.037028 PMID:19703844

Public Health England (2016). NHS Cervical Screening Programme: colposcopy and programme management. 3rd ed. London, UK: Public Health England. Available from: https://www.gov.uk/government/publications/cervical-screening-programme-and-colposcopy-management.

Public Health England (2019a). Cervical screening: guidance for laboratories providing HPV testing and cytology services in the NHS Cervical Screening Programme. Available from: https://www.gov.uk/government/publications/cervical-screening-laboratory-hpv-testing-and-cytology-services.

Public Health England (2019b). NHS Cervical Screening Programme (CSP): liquid-based cytology system conversion training. Available from: https://www.gov.uk/government/publications/

cervical-screening-training-to-use-a-new-cytology-system.

Public Health England (2020). Education pathway. Available from: https://www.gov.uk/government/publications/cervical-screening-cervical-sample-taker-training.

Rebolj M, Rask J, van Ballegooijen M, Kirschner B, Rozemeijer K, Bonde J, et al. (2015). Cervical histology after routine ThinPrep or SurePath liquid-based cytology and computer-assisted reading in Denmark. *Br J Cancer*. 113(9):1259–74. doi:10.1038/bjc.2015.339 PMID:26448176

Rebolj M, van Ballegooijen M, Lynge E, Looman C, Essink-Bot ML, Boer R, et al. (2009). Incidence of cervical cancer after several negative smear results by age 50: prospective observational study. *BMJ*. 338:b1354. doi:10.1136/bmj.b1354 PMID:19395420

Reis NVS, Andrade BB, Guerra MR, Teixeira MTB, Malta DC, Passos VMA (2020). The Global Burden of Disease Study estimates of Brazil's cervical cancer burden. *Ann Glob Health*. 86(1):56. doi:10.5334/aogh.2756 PMID:32566484

Renshaw AA, Elsheikh TM (2013). Assessment of manual workload limits in gynecologic cytology: reconciling data from 3 major prospective trials of automated screening devices. *Am J Clin Pathol*. 139(4):428–33. doi:10.1309/AJCP23YDFKKQNUAO PMID:23525612

Roelens J, Reuschenbach M, von Knebel-Doeberitz M, Wentzensen N, Bergeron C, Arbyn M (2012). $p16^{INK4a}$ immunocytochemistry versus HPV testing for triage of women with minor cytological abnormalities: a systematic review and meta-analysis. *Cancer Cytopathol*. 120(5):294–307. doi:10.1002/cncy.21205 PMID:22700382

Rogovskaya SI, Shabalova IP, Mikheeva IV, Minkina GN, Podzolkova NM, Shipulina OY, et al. (2013). Human papillomavirus prevalence and type-distribution, cervical cancer screening practices and current status of vaccination implementation in Russian Federation, the western countries of the former Soviet Union, Caucasus region and Central Asia. *Vaccine*. 31(Suppl 7): H46–58. doi:10.1016/j.vaccine.2013.06.043 PMID: 24332297

Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Dalla Palma P, Naldoni C, et al. (2007). Accuracy of liquid based versus conventional cytology: overall results of New Technologies for Cervical Cancer Screening randomised controlled trial. *BMJ*. 335(7609):28. doi:10.1136/bmj.39196.740995.BE PMID:17517761

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al.; New Technologies for Cervical Cancer Screening Working Group (2008). 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*. 100(7):492–501. doi:10.1093/jnci/djn065 PMID:18364502

Ronco G, Giorgi-Rossi P, Carozzi F, Dalla Palma P, Del Mistro A, De Marco L, et al.; New Technologies for Cervical Cancer Screening Working Group (2006). Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. *Lancet Oncol*. 7(7):547–55. doi:10.1016/S1470-2045(06)70731-8 PMID:16814206

Ronco G, Montanari G, Confortini M, Parisio F, Berardengo E, Delpiano AM, et al. (2003). Effect of circulation and discussion of cervical smears on agreement between laboratories. *Cytopathology*. 14(3):115–20. doi:10.1046/j.1365-2303.2003.00035.x PMID:12828719

Ronco G, van Ballegooijen M, Becker N, Chil A, Fender M, Giubilato P, et al. (2009). Process performance of cervical screening programmes in Europe. *Eur J Cancer*. 45(15):2659–70. doi:10.1016/j.ejca.2009.07.022 PMID:19713100

Rosenblatt KA, Osterbur EF, Douglas JA (2016). Case-control study of cervical cancer and gynecologic screening: a SEER-Medicare analysis. *Gynecol Oncol*. 142(3):395–400. doi:10.1016/j.ygyno.2016.06.016 PMID:27388696

Rozemeijer K, Naber SK, Penning C, Overbeek LIH, Looman CWN, de Kok IMCM, et al. (2017). Cervical cancer incidence after normal cytological sample in routine screening using SurePath, ThinPrep, and conventional cytology: population based study. *BMJ*. 356:j504. doi:10.1136/bmj.j504 PMID:28196844

Rozemeijer K, Penning C, Siebers AG, Naber SK, Matthijsse SM, van Ballegooijen M, et al. (2016). Comparing SurePath, ThinPrep, and conventional cytology as primary test method: SurePath is associated with increased CIN II^{+} detection rates. *Cancer Causes Control*. 27(1):15–25. doi:10.1007/s10552-015-0678-1 PMID:26458884

Rustagi AS, Kamineni A, Weinmann S, Reed SD, Newcomb P, Weiss NS (2014). Cervical screening and cervical cancer death among older women: a population-based, case-control study. *Am J Epidemiol*. 179(9):1107–14. doi:10.1093/aje/kwu035 PMID:24685531

Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS, et al.; Osmanabad District Cervical Screening Study Group (2005). A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. *Int J Cancer*. 116(4):617–23. doi:10.1002/ijc.21050 PMID:15818610

Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. (2009). HPV screening for cervical cancer in rural India. *N Engl J Med*. 360(14):1385–94. doi:10.1056/NEJMoa0808516 PMID:19339719

Sasieni P, Adams J, Cuzick J (2003). Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. *Br J Cancer*. 89(1):88–93. doi:10.1038/sj.bjc.6600974 PMID:12838306

Sasieni P, Castanon A, Cuzick J (2009). Screening and adenocarcinoma of the cervix. *Int J Cancer*. 125(3):525–9. doi:10.1002/ijc.24410 PMID:19449379

Sasieni PD, Cuzick J, Lynch-Farmery E; The National Co-ordinating Network for Cervical Screening Working Group (1996). Estimating the efficacy of screening by auditing smear histories of women with and without cervical cancer. *Br J Cancer*. 73(8):1001–5. doi:10.1038/bjc.1996.196 PMID:8611418

Schmidt D, Bergeron C, Denton KJ, Ridder R; European CINtec Cytology Study Group (2011). p16/Ki-67 dual-stain cytology in the triage of ASCUS and LSIL Papanicolaou cytology: results from the European Equivocal or Mildly Abnormal Papanicolaou Cytology study. *Cancer Cytopathol*. 119(3):158–66. doi:10.1002/cncy.20140 PMID:21442767

Schneider A, Hoyer H, Lotz B, Leistritza S, Kühne-Heid R, Nindl I, et al. (2000). Screening for high-grade cervical intra-epithelial neoplasia and cancer by testing for high-risk HPV, routine cytology or colposcopy. *Int J Cancer*. 89(6):529–34. doi:10.1002/1097-0215(20001120)89:6<529::AID-IJC11>3.0.CO;2-G PMID:11102899

Seehusen DA, Johnson DR, Earwood JS, Sethuraman SN, Cornali J, Gillespie K, et al. (2006). Improving women's experience during speculum examinations at routine gynaecological visits: randomised clinical trial. *BMJ*. 333(7560):171. doi:10.1136/bmj.38888.588519.55 PMID:16803941

Sepúlveda C, Prado R (2005). Effective cervical cytology screening programmes in middle-income countries: the Chilean experience. *Cancer Detect Prev*. 29(5):405–11. doi:10.1016/j.cdp.2005.07.001 PMID:16188399

Sharp L, Cotton SC, Cruickshank ME, Gray NM, Neal K, Rothnie K, et al.; TOMBOLA Group (2015). Long-term worries after colposcopy: which women are at increased risk? *Womens Health Issues*. 25(5):517–27. doi:10.1016/j.whi.2015.04.002 PMID:26051021

Sherman ME, Dasgupta A, Schiffman M, Nayar R, Solomon D (2007). The Bethesda Interobserver Reproducibility Study (BIRST): a web-based assessment of the Bethesda 2001 System for classifying cervical cytology. *Cancer*. 111(1):15–25. doi:10.1002/cncr.22423 PMID:17186503

Siebers AG, Klinkhamer PJJM, Arbyn M, Raifu AO, Massuger LFAG, Bulten J (2008). Cytologic detection of cervical abnormalities using liquid-based compared with conventional cytology: a randomized controlled trial. *Obstet Gynecol*. 112(6):1327–34. doi:10.1097/AOG.0b013e31818c2b20 PMID:19037043

Siebers AG, Klinkhamer PJJM, Grefte JMM, Massuger LFAG, Vedder JEM, Beijers-Broos A, et al. (2009). Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: a randomized controlled trial. [Erratum in: JAMA. 2009 302(21):2322] *JAMA*. 302(16):1757–64. doi:10.1001/jama.2009.1569 PMID:19861667

Sigurdsson K (2013). Is a liquid-based cytology more sensitive than a conventional Pap smear? *Cytopathology*. 24(4):254–63. doi:10.1111/cyt.12037 PMID:23331613

Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al.; Forum Group Members; Bethesda 2001 Workshop (2002). The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA*. 287(16):2114–9. doi:10.1001/jama.287.16.2114 PMID:11966386

Solomon D, Schiffman M, Tarone R; ALTS Group (2001). Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. *J Natl Cancer Inst*. 93(4):293–9. doi:10.1093/jnci/93.4.293 PMID:11181776

Somrak TM, Sorensen K, Abdul-Karim F (1990). Pap smear: collection, handling and quality assurance. Chicago (IL), USA: ASCP Press.

Spriggs AI (1977). History of cytodiagnosis. *J Clin Pathol*. 30(12):1091–102. doi:10.1136/jcp.30.12.1091 PMID:604355

Sriplung H, Singkham P, Iamsirithaworn S, Jiraphongsa C, Bilheem S (2014). Success of a cervical cancer screening program: trends in incidence in Songkhla, southern Thailand, 1989–2010, and prediction of future incidences to 2030. *Asian Pac J Cancer Prev*. 15(22):10003–8. doi:10.7314/APJCP.2014.15.22.10003 PMID:25520060

Stanczuk GA, Baxter GJ, Currie H, Forson W, Lawrence JR, Cuschieri K, et al. (2017). Defining optimal triage strategies for hrHPV screen-positive women – an evaluation of HPV 16/18 genotyping, cytology, and p16/Ki-67 cytoimmunochemistry. *Cancer Epidemiol Biomarkers Prev*. 26(11):1629–35. doi:10.1158/1055-9965.EPI-17-0534 PMID:28887297

Stoler MH, Baker E, Boyle S, Aslam S, Ridder R, Huh WK, et al. (2020). Approaches to triage optimization in HPV primary screening: extended genotyping and p16/Ki-67 dual-stained cytology – retrospective insights from ATHENA. *Int J Cancer*. 146(9):2599–607. doi:10.1002/ijc.32669 PMID:31490545

Stoler MH, Schiffman M; Atypical Squamous Cells of Undetermined Significance–Low-grade Squamous Intraepithelial Lesion Triage Study (ALTS) Group (2001). Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA*. 285(11):1500–5. doi:10.1001/jama.285.11.1500 PMID:11255427

Strander B, Andersson-Ellström A, Milsom I, Rådberg T, Ryd W (2007). Liquid-based cytology versus conventional Papanicolaou smear in an organized screening program: a prospective randomized study. *Cancer*. 111(5):285–91. doi:10.1002/cncr.22953 PMID:17724676

Sulik SM, Kroeger K, Schultz JK, Brown JL, Becker LA, Grant WD (2001). Are fluid-based cytologies superior to the conventional Papanicolaou test? A systematic review. *J Fam Pract*. 50(12):1040–6. PMID:11742605

Sultana F, English DR, Simpson JA, Canfell K, Gertig DM, Saville M (2014). High-grade cervical abnormalities and cervical cancer in women following a negative Pap smear with and without an endocervical component: a cohort study with 10 years of follow-up. *Int J Cancer*. 135(5):1213–9. doi:10.1002/ijc.28756 PMID:24488882

Swailes AL, Hossler CE, Kesterson JP (2019). Pathway to the Papanicolaou smear: the development of cervical cytology in twentieth-century America and implications in the present day. *Gynecol Oncol*. 154(1):3–7. doi:10.1016/j.ygyno.2019.04.004 PMID:30995961

Sykes PH, Harker DY, Miller A, Whitehead M, Neal H, Wells JE, et al. (2008). A randomised comparison of SurePath liquid-based cytology and conventional smear cytology in a colposcopy clinic setting. *BJOG*. 115(11):1375–81. doi:10.1111/j.1471-0528.2008.01865.x PMID:18823488

Talbott EO, Norman SA, Kuller LH, Ishii EK, Baffone KM, Dunn MS, et al. (1995). Refining preventive strategies for invasive cervical cancer: a population-based case-control study. *J Womens Health*. 4(4):387–95. doi:10.1089/jwh.1995.4.387

Tanabodee J, Thepsuwan K, Karalak A, Laoaree O, Krachang A, Manmatt K, et al. (2015). Comparison of efficacy in abnormal cervical cell detection between liquid-based cytology and conventional cytology. *Asian Pac J Cancer Prev*. 16(16):7381–4. doi:10.7314/APJCP.2015.16.16.7381 PMID:26514540

Tarkkanen J, Geagea A, Nieminen P, Anttila A (2003). Quality improvement project in cervical cancer screening: practical measures for monitoring laboratory performance. *Acta Obstet Gynecol Scand*. 82(1):82–8. doi:10.1034/j.1600-0412.2003.820115.x PMID:12580846

Taylor S, Kuhn L, Dupree W, Denny L, De Souza M, Wright TC Jr (2006). Direct comparison of liquid-based and conventional cytology in a South African screening trial. *Int J Cancer*. 118(4):957–62. doi:10.1002/ijc.21434 PMID:16152600

Taylor VM, Yasui Y, Burke N, Nguyen T, Acorda E, Thai H, et al. (2004). Pap testing adherence among Vietnamese American women. *Cancer Epidemiol Biomarkers Prev*. 13(4):613–9. PMID:15066927

Thangarajah F, Einzmann T, Bergauer F, Patzke J, Schmidt-Petruschkat S, Theune M, et al. (2016). Cervical screening program and the psychological impact of an abnormal Pap smear: a self-assessment questionnaire study of 590 patients. *Arch Gynecol Obstet*. 293(2):391–8. doi:10.1007/s00404-015-3821-z PMID:26202136

Theil KS (2012). Bone marrow processing and normal morphology. In: Kottke-Marchant K, Davis BH, editors. Laboratory hematology practice. 1st ed. Oxford, UK: Wiley-Blackwell; pp. 281–99. doi:10.1002/9781444398595.ch22

Thrall MJ (2019). Automated screening of Papanicolaou tests: a review of the literature. *Diagn Cytopathol*. 47(1):20–7. doi:10.1002/dc.23931 PMID:29603675

Tsoumpou I, Arbyn M, Kyrgiou M, Wentzensen N, Koliopoulos G, Martin-Hirsch P, et al. (2009). p16^{INK4a} immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev*. 35(3):210–20. doi:10.1016/j.ctrv.2008.10.005 PMID:19261387

UNFPA (2019). UNFPA Tajikistan. Key results of Tajikistan in 2019. United Nations Population Fund. Available from: https://www.unfpa.org/data/transparency-portal/unfpa-tajikistan.

Vaccarella S, Franceschi S, Engholm G, Lönnberg S, Khan S, Bray F (2014). 50 years of screening in the Nordic countries: quantifying the effects on cervical cancer incidence. *Br J Cancer*. 111(5):965–9. doi:10.1038/bjc.2014.362 PMID:24992581

Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F (2013). Worldwide trends in cervical cancer incidence: impact of screening against changes in disease risk factors. *Eur J Cancer*. 49(15):3262–73. doi:10.1016/j.ejca.2013.04.024 PMID:23751569

van Ballegooijen M, Hermens R (2000). Cervical cancer screening in the Netherlands. *Eur J Cancer*. 36(17):2244–6. doi:10.1016/S0959-8049(00)00317-8 PMID:11072213

Vicus D, Sutradhar R, Lu Y, Elit L, Kupets R, Paszat L; Investigators of the Ontario Cancer Screening Research Network (2014). The association between cervical cancer screening and mortality from cervical cancer: a population based case-control study. *Gynecol Oncol*. 133(2):167–71. doi:10.1016/j.ygyno.2014.02.037 PMID:24589414

Vicus D, Sutradhar R, Lu Y, Kupets R, Paszat L; Ontario Cancer Screening Research Network (2015). Association between cervical screening and prevention of invasive cervical cancer in Ontario: a population-based case-control study. *Int J Gynecol Cancer*. 25(1):106–11. doi:10.1097/IGC.0000000000000305 PMID:25377725

Virani S, Sriplung H, Bilheem S, Sripan P, Maneesai P, Waisri N, et al. (2018). Effect of the national screening program on malignancy status of cervical cancer in northern Thailand. *Int J Public Health*. 63(3):377–85. doi:10.1007/s00038-018-1077-7 PMID:29362857

von Knebel Doeberitz M (2002). New markers for cervical dysplasia to visualise the genomic chaos created by aberrant oncogenic papillomavirus infections. *Eur J Cancer*.

38(17):2229–42. doi:10.1016/S0959-8049(02)00462-8 PMID:12441259

von Knebel Doeberitz M, Rittmüller C, zur Hausen H, Dürst M (1992). Inhibition of tumorigenicity of cervical cancer cells in nude mice by HPV E6-E7 anti-sense RNA. *Int J Cancer*. 51(5):831–4. doi:10.1002/ijc.2910510527 PMID:1319412

Wang J, Andrae B, Sundström K, Ploner A, Ström P, Elfström KM, et al. (2017). Effectiveness of cervical screening after age 60 years according to screening history: nationwide cohort study in Sweden. *PLoS Med*. 14(10):e1002414. doi:10.1371/journal.pmed.1002414 PMID:29065127

Wang J, Elfström KM, Andrae B, Nordqvist Kleppe S, Ploner A, Lei J, et al. (2020). Cervical cancer case–control audit: results from routine evaluation of a nationwide cervical screening program. *Int J Cancer*. 146(5):1230–40. doi:10.1002/ijc.32416 PMID:31107987

Wardle J, Pernet A, Stephens D (1995). Psychological consequences of positive results in cervical cancer screening. *Psychol Health*. 10(3):185–94. doi:10.1080/08870449508401948

Weiss VL, Cate F, Bloom L, Fadare O, Coogan AC, Desouki MM (2016). Age cut-off for reporting endometrial cells on a Papanicolaou test: 50 years may be more appropriate than 45 years. *Cytopathology*. 27(4):242–8. doi:10.1111/cyt.12280 PMID:26689647

Wentzensen N, Bergeron C, Cas F, Eschenbach D, Vinokurova S, von Knebel Doeberitz M (2005). Evaluation of a nuclear score for $p16^{INK4a}$-stained cervical squamous cells in liquid-based cytology samples. *Cancer*. 105(6):461–7. doi:10.1002/cncr.21378 PMID:16116604

Wentzensen N, Clarke MA, Bremer R, Poitras N, Tokugawa D, Goldhoff PE, et al. (2019). Clinical evaluation of human papillomavirus screening with p16/Ki-67 dual stain triage in a large organized cervical cancer screening program. *JAMA Intern Med*. 179(7):881–8. doi:10.1001/jamainternmed.2019.0306 PMID:31081870

Wentzensen N, Fetterman B, Castle PE, Schiffman M, Wood SN, Stiemerling E, et al. (2015). p16/Ki-67 dual stain cytology for detection of cervical precancer in HPV-positive women. *J Natl Cancer Inst*. 107(12):djv257. doi:10.1093/jnci/djv257 PMID:26376685

Wentzensen N, Schiffman M, Palmer T, Arbyn M (2016). Triage of HPV positive women in cervical cancer screening. *J Clin Virol*. 76(Suppl 1):S49–55. doi:10.1016/j.jcv.2015.11.015 PMID:26643050

Wentzensen N, Schwartz L, Zuna RE, Smith K, Mathews C, Gold MA, et al. (2012). Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. *Clin Cancer Res*. 18(15):4154–62. doi:10.1158/1078-0432.CCR-12-0270 PMID:22675168

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. [Erratum in: Ann Intern Med. 2012 156(1 Pt 1):71–2] *Ann Intern Med*. 155(10):687–97, W214–5. doi:10.7326/0003-4819-155-10-201111150-00376 PMID:22006930

WHO (2019). Better noncommunicable disease outcomes: challenges and opportunities for health systems. Turkmenistan county assessment. Copenhagen, Denmark: World Health Organization Regional Office for Europe. Available from: https://www.euro.who.int/__data/assets/pdf_file/0013/403015/HSS-NCDs-TKM_WEB-1.pdf.

Wilbur DC, Black-Schaffer WS, Luff RD, Abraham KP, Kemper C, Molina JT, et al. (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(5):767–75. doi:10.1309/AJCP8VE7AWBZCVQT PMID:19846820

Williams ARW (2006). Liquid-based cytology and conventional smears compared over two 12-month periods. *Cytopathology*. 17(2):82–5. doi:10.1111/j.1365-2303.2006.00339.x PMID:16548992

Wright TC Jr, Behrens CM, Ranger-Moore J, Rehm S, Sharma A, Stoler MH, et al. (2017). Triaging HPV-positive women with p16/Ki-67 dual-stained cytology: results from a sub-study nested into the ATHENA trial. *Gynecol Oncol*. 144(1):51–6. doi:10.1016/j.ygyno.2016.10.031 PMID:28094038

Xu L, Verdoodt F, Wentzensen N, Bergeron C, Arbyn M (2016). Triage of ASC-H: a meta-analysis of the accuracy of high-risk HPV testing and other markers to detect cervical precancer. *Cancer Cytopathol*. 124(4):261–72. doi:10.1002/cncy.21661 PMID:26618614

Yabroff KR, Saraiya M, Meissner HI, Haggstrom DA, Wideroff L, Yuan G, et al. (2009). Specialty differences in primary care physician reports of Papanicolaou test screening practices: a national survey, 2006 to 2007. *Ann Intern Med*. 151(9):602–11. doi:10.7326/0003-4819-151-9-200911030-00005 PMID:19884621

Yang B, Morrell S, Zuo Y, Roder D, Tracey E, Jelfs P (2008). A case-control study of the protective benefit of cervical screening against invasive cervical cancer in NSW women. *Cancer Causes Control*. 19(6):569–76. doi:10.1007/s10552-008-9118-9 PMID:18286380

Yang DX, Soulos PR, Davis B, Gross CP, Yu JB (2018). Impact of widespread cervical cancer screening: number of cancers prevented and changes in race-specific incidence. *Am J Clin Oncol*. 41(3):289–94. doi:10.1097/COC.0000000000000264 PMID:26808257

Yanikkerem E, Özdemir M, Bingol H, Tatar A, Karadeniz G (2009). Women's attitudes and expectations regarding gynaecological examination. *Midwifery*. 25(5):500–8. doi:10.1016/j.midw.2007.08.006 PMID:18086509

Yu CK, Rymer J (1998). Women's attitudes to and awareness of smear testing and cervical cancer. *Br J Fam Plann*. 23(4):127–33. PMID:9882766

Zappa M, Visioli CB, Ciatto S, Iossa A, Paci E, Sasieni P (2004). Lower protection of cytological screening for adenocarcinomas and shorter protection for younger women: the results of a case–control study in Florence. *Br J Cancer*. 90(9):1784–6. doi:10.1038/sj.bjc.6601754 PMID:15150597

4.4 HPV testing

4.4.1 Technical descriptions

(a) Introduction

It has long been recognized that there is a strong etiological link between persistent infection with certain HPV types and subsequent development of cervical precancer and cancer. This has led to the idea that the detection of sequences of the HPV genome could become an alternative screening tool that could replace screening by the microscopic examination of cervical cells (IARC, 2005, 2007, 2012; Bouvard et al., 2009; see also Sections 1.2.1 and 1.2.2).

The HPV genome is a circular, double-stranded DNA molecule that codes for two late proteins (L1 and L2), which form the capsid, and several early (E) genes, which code for various proteins that are important for diverse viral functions. The E6 and E7 proteins are essential for the transformation of infected cells towards neoplasia (IARC, 2007, 2012).

Large RCTs have demonstrated that women with a negative hrHPV DNA test result have lower risks of CIN3 and cervical cancer than women with normal cervical cytology; therefore, many countries are moving towards screening with HPV tests (Arbyn et al., 2012; Huh et al., 2015; Machalek et al., 2019; Ronco et al., 2014; von Karsa et al., 2015). Currently, a multitude of hrHPV assays are available, but only a few have been clinically validated for use in cervical cancer screening against internationally agreed clinical criteria (Poljak et al., 2020). This section discusses HPV nucleic acid tests that detect DNA or RNA sequences of alpha HPV types that are considered to be carcinogenic, i.e. the 12 types classified as carcinogenic to humans (Group 1): HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59. HPV68, which is probably carcinogenic to humans (Group 2A), and HPV66, which is possibly carcinogenic to humans (Group 2B), are often included in the panel of types targeted by the hrHPV tests (Bernard et al., 2010; IARC, 2012), although their etiological fraction in cervical cancer carcinogenesis is very low and their inclusion decreases the clinical specificity of such tests (see Sections 1.2.1 and 1.2.2 and Figs. 1.9 and 1.10).

(b) Categories of HPV nucleic acid tests

hrHPV assays can be classified by the following parameters: the nucleic acid targeted (viral genomic DNA [HPV DNA tests] or viral messenger RNA [mRNA] [HPV RNA tests]), the viral genes targeted, the level of genotyping detail, whether signal amplification (e.g. hybrid capture) or target amplification (e.g. polymerase chain reaction [PCR] or next-generation sequencing) is used, the method of identification of amplicons, the output result (qualitative or quantitative), and the inclusion of internal controls that check the validity of the specimen. An inventory of more than 200 HPV tests that were available in 2020 and are classified according to these principles is available in Poljak et al. (2020).

The main applied test systems used to identify HPV nucleic acid sequences are hybridization and PCR. In hybrid capture, RNA probes hybridize with complementary HPV DNA if present in a sample; the DNA/RNA hybrids are subsequently captured by anti-DNA/RNA antibodies coupled to an enzyme that generates a chemical reaction and yields a quantified light signal (Lorincz, 1997). In PCR systems, one or more adjacent pairs of oligonucleotide primers directed to the 3′ and 5′ ends of a target sequence will bind to it and initialize amplification of the DNA between the primers by the temperature-sensitive Taq DNA polymerase. The amplified target DNA is called an amplicon. After multiple cycles of amplification, controlled by alternating the temperature, a large number of amplicons are generated. PCRs targeting short amplicons are analytically more sensitive than those targeting a longer amplicon (Iftner & Villa, 2003). Diverse systems are used to identify the amplicons. In real-time PCR, a

quantified light signal is generated that is correlated with the amount of target DNA (Josefsson et al., 1999). Real-time PCR can also be applied in multiplex format, in which the presence of and viral load of multiple carcinogenic HPV types can be assessed simultaneously and with control of the amount of input DNA (Moberg et al., 2004).

The identification of hrHPV DNA indicates the presence of the virus, whereas the presence of hrHPV RNA may serve as an indication of viral activity, and it has therefore been proposed by some researchers to be a more specific marker of cervical neoplasia than DNA (Haedicke & Iftner, 2016).

HPV tests can target multiple sequences throughout the viral genome or specific parts of a given viral gene. Many tests target the well-conserved part of the *L1* gene, whereas others target *E* genes. Viral integration in the human genome, which often occurs in the *E2* region, results in interruption of HPV DNA and enhanced transcription of the *E6–E7* sequence, which may predispose the cell to neoplastic transformation (zur Hausen, 2002). However, this molecular pathogenetic pathway has been challenged by HPV genome-wide next-generation sequencing analyses, which indicate that integration into the host DNA can occur almost anywhere throughout the viral genome (Hu et al., 2015; Dyer et al., 2016). Moreover, no epidemiological evidence is currently available that indicates differences in diagnostic accuracy between tests targeting different genes (Arbyn et al., 2015).

With regard to the level of detail in HPV genotyping, the following can be distinguished: (i) no genotyping; (ii) limited genotyping, in which the most carcinogenic HPV types, HPV16 or HPV18 with or without HPV45, are distinguished from the other hrHPV types; (iii) extended genotyping, in which more hrHPV types – but not all – are distinguished separately; and (iv) full genotyping assays, which identify all individual hrHPV types of the high-risk group separately. Some full genotyping tests detect additional individual HPV types that do not belong to the high-risk group. Certain types (HPV types 26, 53, 66, 67, 73, and 82) are possibly carcinogenic to humans (Group 2B). Their inclusion in HPV screening tests would increase the number of false-positive results and increase the burden of follow-up, cost, and harms associated with screening (see also Sections 1.2.1 and 1.2.2). Epidemiological research is under way to investigate whether all 12 HPV types classified as carcinogenic to humans (Group 1) should be routinely detected in primary HPV screening in an optimally efficient screening programme.

(c) Clinical applications of HPV testing

HPV tests can be used for several clinical purposes: (1) as a primary cervical cancer screening test, alone or in combination with cytology (co-testing); (2) as a triage test for women with minor abnormal cervical cytology in the context of cytology-based screening; (3) for the triage of women with a positive primary hrHPV screening test result by genotyping, or as delayed triage when the reflex triage test result is negative; and (4) to monitor the success or failure of treatment of a precancerous lesion. Triage of hrHPV-positive women (application 3), distinguishes between (i) reflex triage with genotyping, in which the detection of the most carcinogenic types (HPV16 or HPV18) triggers referral to colposcopy, leaving women who are positive only for other hrHPV types to be triaged further, and (ii) delayed triage of hrHPV-positive women who had a negative reflex HPV triage test result. Reflex triage is the immediate testing with markers using the same specimen used for primary screening. New triage strategies propose to fine-tune the management of hrHPV-positive women according to the risk of present or incipient CIN3+ associated with individual genotypes or groups of genotypes (Cheung et al., 2020; Demarco et al., 2020).

In addition to clinical purposes, HPV tests can also be used for epidemiological research and to evaluate the effects of HPV vaccination. To measure the effects of HPV vaccination in trials, high analytical sensitivity is required, whereas in clinical applications accuracy for clinically relevant outcomes is important (as discussed further below) (WHO, 2010; Dillner et al., 2011). High-grade cervical lesions including CIN2+ (in particular, CIN3+) and AIS, and cervical SCC and adenocarcinoma of the cervix are all relevant clinical outcomes (Herbert et al., 2008).

HPV tests are typically performed on cervical specimens taken by health-care workers, but they can also be performed on self-collected vaginal samples or urine and on tissue specimens. This section focuses on the use of HPV tests in cervical cancer screening using cervical samples taken by a health professional. The use of HPV testing in other settings is described elsewhere: HPV genotyping in triage of hrHPV-positive women in Section 4.4.7 and hrHPV testing on self-collected samples and the use of HPV RNA testing in Sections 4.4.5 and 4.4.6, respectively.

In primary screening, hrHPV tests should yield results that are informative about the risk of having or developing cervical precancer or cancer and should have a balanced clinical sensitivity and specificity. Infections with low concentrations of virus, in particular infections with less carcinogenic hrHPV types that usually clear spontaneously, should ideally not be detected by a screening test (Snijders et al., 2003; Eklund et al., 2014).

(i) Principles of HPV test validation

In 2009, an international team of virologists and clinical epidemiologists defined the minimum requirements that HPV assays should fulfil for them to be accepted for use in cervical cancer screening (Meijer et al., 2009). Two tests were accepted as standard comparator tests: Hybrid Capture 2 (HC2) and GP5+/6+ PCR enzyme immunoassay (EIA). Four large population-based RCTs, conducted in Europe, have provided consistent evidence that screening with these assays provides better protection against future CIN3 or cancer compared with good-quality cytology (Arbyn et al., 2012; Ronco et al., 2014). However, to validate other hrHPV DNA assays, it is not required to set up RCTs with long-term follow-up. It is deemed sufficient that three criteria (Table 4.21) are fulfilled to accept another hrHPV DNA test for use in primary cervical cancer screening. The given hrHPV DNA test (the index test) should have non-inferior cross-sectional sensitivity and specificity for CIN2+ compared with one of the comparator assays (HC2 or GP5+/6+ PCR EIA) (Meijer et al., 2009). The agreed benchmarks (index test divided by standard comparator test) are 0.90 for relative sensitivity and 0.98 for relative specificity. The paired statistical test for non-inferiority will be significant when the lower bound of the 90% confidence interval around the relative sensitivity or relative specificity is greater than or equal to the benchmark (Tang et al., 2003). A representative set of cervical samples (at least 60 CIN2+ cases and at least 800 < CIN2 cases) derived from a population-based screening cohort should be selected (Meijer et al., 2009). Moreover, the new test should show high intralaboratory and interlaboratory reproducibility, with a lower bound of the 95% confidence interval of at least 87% or a kappa of at least 0.5 (Meijer et al., 2009). The recommended sample size for the reproducibility assessment is at least 500 with an hrHPV prevalence of 30% as established with a standard comparator test (Table 4.21). These guidelines apply only to hrHPV DNA testing. For screening tests using targets other than hrHPV DNA (e.g. HPV RNA, methylation markers, protein markers, or other test systems), additional longitudinal criteria are needed. For HPV DNA tests, these longitudinal data are not needed because the longitudinal safety (low 5-year risk of cancer after an earlier negative test result) is established through RCTs and supported by observational

Table 4.21 International validation criteria for high-risk human papillomavirus (hrHPV) DNA tests acceptable for use in primary cervical cancer screening, based on the relative accuracy for detection of cervical intraepithelial neoplasia grade 2 or worse (CIN2+) of an index HPV test compared with a standard comparator test[a]

Criteria	Study population needed	Target
1. Relative sensitivity[b]	≥ 60 samples from women with CIN2+	*P* for non-inferiority < 0.05[c] (accepting 0.90 as benchmark) The lower bound of the 90% CI should be ≥ 0.90
2. Relative specificity[b]	≥ 800 samples from women with < CIN2	*P* for non-inferiority < 0.05[c] (accepting 0.98 as benchmark) The lower bound of the 90% CI should be ≥ 0.98
3. Intralaboratory and interlaboratory reproducibility	≥ 500 samples from a screening population with an hrHPV prevalence of 30% (as established with a standard comparator test)	Lower bound of the 95% CI ≥ 87% Kappa ≥ 0.5

CI, confidence interval; CIN2, cervical intraepithelial neoplasia grade 2; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; hrHPV, high-risk human papillomavirus.

[a] Standard comparator tests: Hybrid Capture 2 and GP5+/6+ polymerase chain reaction (PCR) enzyme immunoassay (EIA). These two tests have been validated through randomized controlled trials that demonstrated lower incidence of cervical cancer compared with good-quality cytology.

[b] Relative accuracy of the index hrHPV DNA test compared with the standard comparator test for the outcome CIN2+.

[c] One-sided non-inferiority test for paired data accepting a power of 90% and a confidence level of 95% (Tang et al., 2003). Because this statistical test is one-sided, the equivalent confidence level for the lower bound of the CI (two-sided expression) should be 90%.

Compiled from Meijer et al. (2009).

longitudinal studies. However, for other molecular targets, a high cross-sectional sensitivity does not provide sufficient evidence that the lead-time gain (time span between detectability of a neoplastic lesion and when it becomes clinically manifest) is similar to that for HPV DNA and that use of the same screening interval as that proposed for hrHPV DNA screening tests (usually 5 years or longer) can be accepted as safe.

(ii) Updating and extension of HPV test validation guidelines

The international validation criteria (Meijer et al., 2009) are for hrHPV DNA testing on cervical samples. Currently, new criteria are being developed that will include HPV genotyping and HPV testing on alternative specimens (self-collected vaginal samples or urine) and may involve standard comparator tests other than HC2 and GP5+/6+ PCR EIA (Arbyn & Hillemanns, 2018). Recent meta-analyses indicated that HPV tests based on a principle of signal amplification (e.g. HC2 or *care*HPV) are less sensitive and specific for the detection of CIN2+ on self-collected vaginal samples than on clinician-collected cervical samples. RNA-based HPV assays are less sensitive on self-collected samples. However, PCR-based hrHPV DNA assays, validated on cervical specimens, seem to be as sensitive and nearly as specific on vaginal samples as they are on cervical samples (Arbyn et al., 2014, 2018).

(iii) Assays that detect molecules other than hrHPV DNA

An HPV RNA assay targeting E6/E7 transcripts of only five HPV types (HPV types 16, 18, 31, 33, and 45) was significantly less sensitive but more specific than the standard comparator hrHPV DNA tests (Arbyn et al., 2015). Another RNA HPV assay targeting E6/E7 transcripts of 14 hrHPV types in bulk fulfils the three international cross-sectional validation criteria described in Table 4.21 (Arbyn et al., 2015). The assessment of its longitudinal performance and risk of CIN3+ after baseline testing with an RNA

test versus after testing with a validated DNA test is covered in Section 4.4.6.

(iv) Other important factors that influence the choice of a screening test

In addition to accuracy, other characteristics need to be taken into account when choosing a screening test, such as the availability of the assay, reagents, and disposables, the throughput capacity and turnaround time (time span between arrival of the specimen and communication of the result), costs, applicability on samples taken by the woman (self-collected vaginal samples or urine), the requirement for equipped laboratories, user-friendliness, the need for running water and electricity, the possibility of point-of-care testing, and the possibility of providing triage information (genotyping or viral load). A comprehensive overview of logistic, regulatory, managerial, training, and quality control aspects of the choice of HPV assays, procurement, sample collection, transport of specimens to the laboratory, pre-analytical handling, testing, and result communication was given in a recent WHO document (WHO, 2020a).

Most of the assays that have been validated to date for screening require a well-equipped laboratory to perform the HPV tests. Two hrHPV DNA assays, one using the hybrid capture principle and the other using a cartridge, are prequalified by WHO for hrHPV testing in field conditions in low-resource countries (WHO, 2019). Point-of-care hrHPV testing is particularly relevant for screen-and-treat strategies (see Section 5.1).

4.4.2 Comparison of HPV DNA testing versus cytology

(a) Introduction

The evidence for HPV DNA testing as a modality for primary cervical screening has been accumulating for two decades. From first principles, molecular testing for the presence of HPV provides a sensitive assessment of a woman's risk of currently harbouring, or in the future developing, a precancer or invasive cervical cancer, because nearly all cervical cancers are caused by HPV infection.

In the 2005 *IARC Handbook* on cervical cancer screening (IARC, 2005), the performance of HPV assays in the detection of precancerous lesions was compared with that of cytology. At the time, almost all of the evidence was from cross-sectional studies, and there was no prospective evaluation of the impact of primary HPV screening on invasive cervical cancer. Nevertheless, the *Handbook* concluded: "For primary screening of women older than 30 years of age, HPV testing yields on average about 10–20% greater sensitivity and 10% lower specificity than cytology (either conventional or liquid-based). In some studies, the combination of cytology and HPV testing (as independent or reflex testing) attained very high sensitivity and negative predictive values (approaching 100%). A testing combination with such a high negative predictive value could potentially allow screening intervals to be increased, e.g., from the minimum of three years up to five years or longer, depending on the population and risk profile. The drawback of this approach is the loss in specificity with respect to either test in isolation due to the excessive number of patients who would need to be referred for colposcopy."

Since the publication of the 2005 *IARC Handbook*, the evidence base on the sensitivity and NPV of HPV DNA testing versus cytology has become substantially larger, and direct evidence has become available on the protection provided by HPV-based and cytology-based screening against cervical cancer and death from cervical cancer. Furthermore, the screening process for CIN2+ and CIN3+ has been evaluated in the context of a combination of measures taken to increase specificity and minimize harms, including the appropriate use of triage of HPV-positive women (see Section 4.4.7 and

Section 4.4.8). The evidence base for the relative performance of HPV and cytology screening now includes: (i) cross-sectional diagnostic studies, which have been synthesized in meta-analyses to provide evidence on the relative sensitivity and specificity of HPV DNA testing versus cytology for the detection of CIN2 and CIN3; (ii) evidence from longitudinal RCTs, mainly in high-income countries, to evaluate whether the increased detection of CIN2+ with HPV testing results in a decrease in CIN2+ in the subsequent screening round; (iii) evidence from a major RCT of HPV DNA testing versus cytology versus VIA screening in India, with cervical cancer incidence and mortality outcomes, and evidence from individual data of four RCTs in Europe that were pooled to evaluate the effect on cancer incidence; (iv) randomized health services trials and national, regional, and pilot screening programmes, which provide information about the impact of HPV-based screening, sometimes with new, less-aggressive protocols, on the detection of CIN3+ and on resource consumption, and which will provide evidence about effectiveness, and (v) longitudinal studies of women screened by HPV testing and cytology, which are particularly relevant for defining risk-based screening intervals.

This experience, combined with well-validated modelling of the longer-term effects of scaled-up HPV testing, has supported the increased use of HPV testing as the sole primary screening test (or, in a few settings, as a co-test with cytology) in high-income countries and the recommendation to support HPV testing in the 2020 WHO strategic plan for the elimination of cervical cancer as a public health problem (WHO, 2020b). Since 2017, several high-income countries have transitioned from cytology screening to primary HPV screening programmes at screening intervals of 5 years or longer, and this is increasingly also providing evidence on the real-world experience with HPV screening.

(b) Diagnostic studies

A Cochrane review published in 2017 compared the accuracy of HPV testing and cervical cytology for the detection of CIN2+ and CIN3+ in women who were participating in cervical cancer screening and who were not being followed up for previous cytological abnormalities (Koliopoulos et al., 2017). This systematic review and meta-analysis searched for articles published between 1992 and 2015. The review focused on studies in which all women received both HPV testing and cervical cytology. A combination of colposcopy and histology was used as the reference standard. If at least one of the screening tests was positive, women underwent colposcopy with directed biopsy of abnormal areas and histological verification. Women did not know their disease status at the time of recruitment. Of the 40 eligible studies, which included more than 140 000 women, 29 studies conducted head-to-head comparison of HPV DNA testing by signal amplification or target amplification versus conventional cytology or LBC (Pap) testing using a threshold of ASC-US for the detection of CIN2+ or CIN3+.

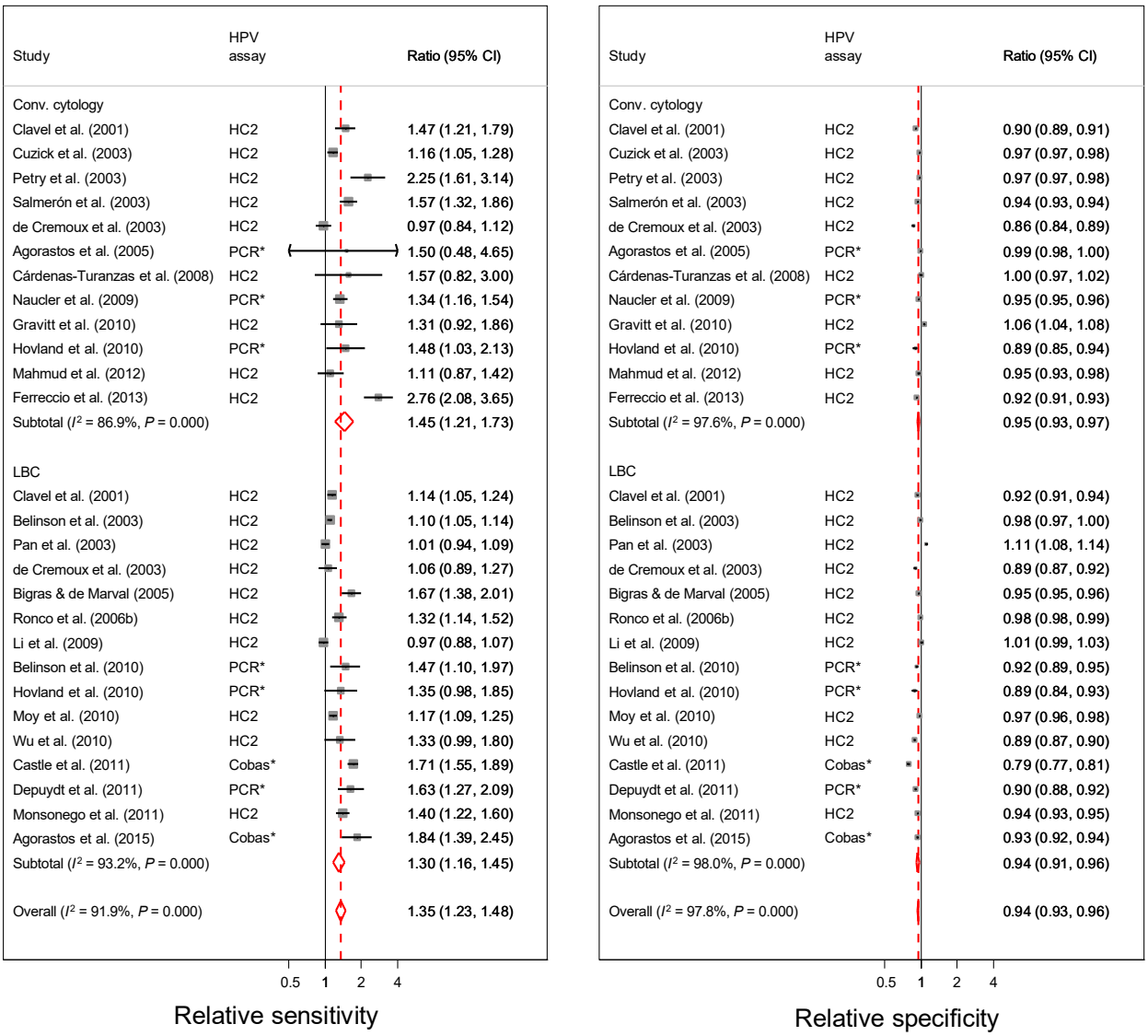
For the detection of CIN2+, the sensitivity of HPV DNA-based tests was higher than that of cytology methods (pooled relative sensitivity, 1.35; 95% CI, 1.23–1.48) and the specificity was lower (pooled relative specificity, 0.94; 95% CI, 0.93–0.96) (Fig. 4.2). For the detection of CIN3+, the pooled relative sensitivity was 1.37 (95% CI, 1.20–1.55) and the pooled relative specificity was 0.95 (95% CI, 0.94–0.97) (Fig. 4.3).

(c) RCTs

(i) Description

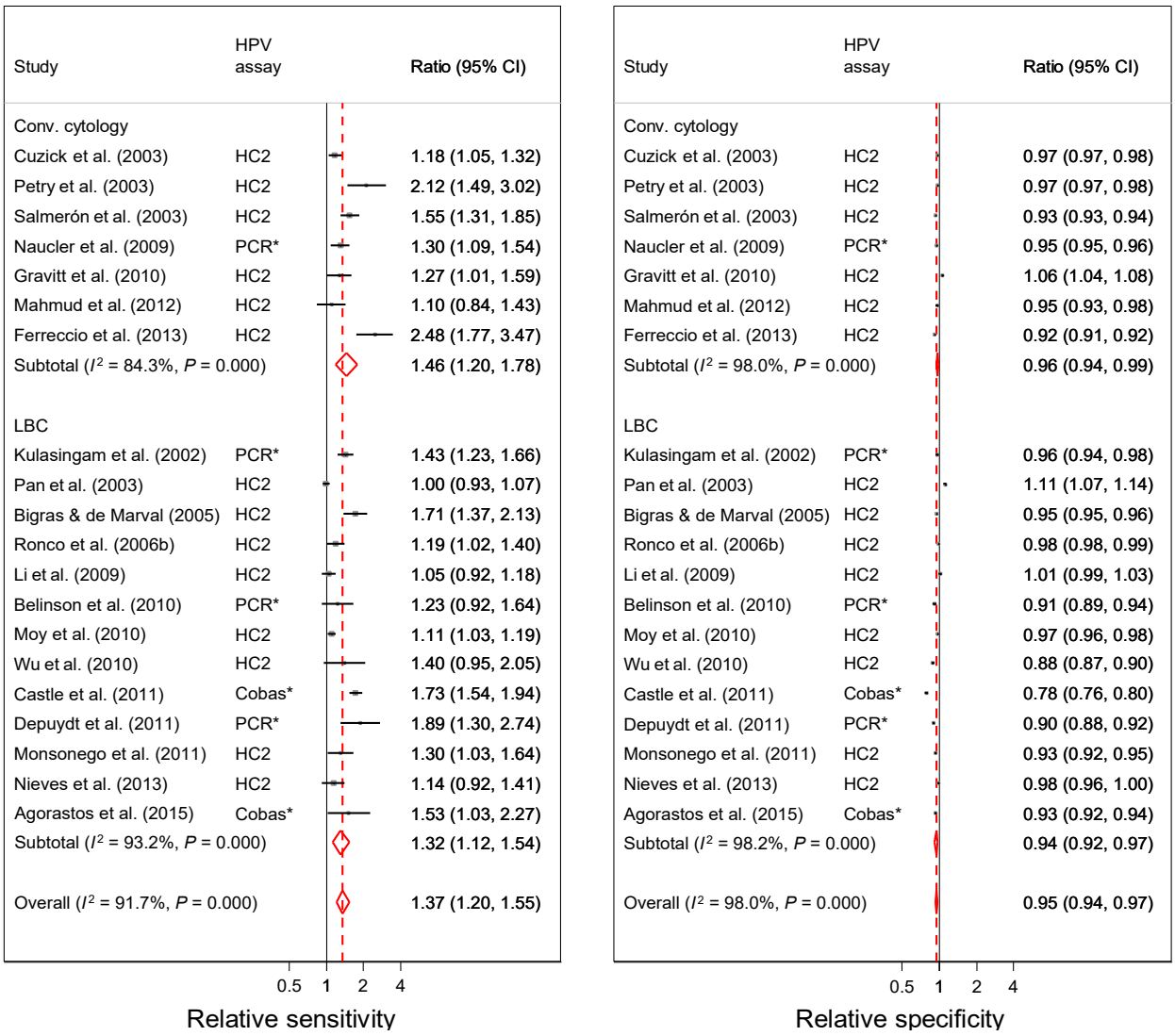
When the 2005 *IARC Handbook* was published, large RCTs of HPV testing in primary cervical cancer screening were in progress but had not yet reported longitudinal outcomes. Since then, eight major RCTs comparing HPV DNA-based screening with cytology-based

Fig. 4.2 Relative sensitivity (left) and relative specificity (right) of hrHPV testing compared with cytology at a threshold of ASC-US+ for the detection of CIN2+



ASC-US+, atypical squamous cells of undetermined significance or worse; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; Cobas*, cobas 4800; Conv., conventional; HC2, Hybrid Capture 2; hrHPV, high-risk human papillomavirus; LBC, liquid-based cytology; PCR*, polymerase chain reaction-based assay targeting at least 13 carcinogenic HPV types.
Created by the Working Group with data from Koliopoulos et al. (2017).

Fig. 4.3 Relative sensitivity (left) and relative specificity (right) of hrHPV testing compared with cytology at a threshold of ASC-US+ for the detection of CIN3+



ASC-US+, atypical squamous cells of undetermined significance or worse; CI, confidence interval; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; Cobas*, cobas 4800; Conv., conventional; HC2, Hybrid Capture 2; hrHPV, high-risk human papillomavirus; LBC, liquid-based cytology; PCR*, polymerase chain reaction-based assay targeting at least 13 carcinogenic HPV types.
Created by the Working Group with data from Koliopoulos et al. (2017).

screening have reported results. An important goal of the RCTs was to evaluate whether the excess CIN2+ detected by HPV DNA-based screening represented clinically relevant persistent disease. For this purpose, women were randomly assigned to HPV DNA-based testing or cytology-based screening at enrolment, and it was investigated whether an increase in detection of CIN2+ in the intervention arm versus the control arm in the first round was followed by a decrease in the second round. In addition, to avoid bias, in the second round in most studies the same screening methodology was applied in both arms. RCTs have also been used to study the benefits of combined HPV DNA testing and cytology (co-testing) compared with primary HPV DNA testing. Those analyses are reviewed in Section 4.4.4. Brief descriptions of the characteristics of the eight major RCTs are given here.

Five RCTs were conducted in European countries, all within organized screening programmes in which the target population was actively invited to primary screening and, if needed, triage testing and treatment. These programmes routinely recorded the numbers of women invited, screened, and treated.

The New Technologies for Cervical Cancer Screening (NTCC) trial was conducted at nine participating centres in Italy and enrolled a total of 94 370 women aged 25–60 years over two implementation phases in 2002–2004. In the intervention arm, co-testing with HPV (HC2) testing and LBC was applied in the first phase (45 174 women enrolled in 2002–2003) and stand-alone HPV testing was applied in the second phase (49 196 women enrolled in 2002–2004). In the first phase, participants in the intervention arm younger than 35 years were referred for colposcopy if they were ASC-US+ or if they were HPV-positive and/or ASC-US+ after 1 year. Women aged 35 years and older were referred for colposcopy if they were HPV-positive and/or ASC-US+. In the second phase, all HPV-positive women were immediately referred for colposcopy, irrespective of age. In the control arm, women were screened using conventional cytology alone. In the second round, all women were screened using conventional cytology, and no further HPV testing was done. Results from the first two rounds of screening, with a 3-year interval (total follow-up period, 7 years), have been published (Ronco et al., 2006a, b, 2008, 2010).

The Population Based Screening Study Amsterdam (POBASCAM) trial was conducted in the Greater Amsterdam region in the Netherlands. Women aged 29–61 years were recruited in 1999–2002. A total of 44 102 women were enrolled and randomized either to co-testing with HPV DNA (GP5+/6+ PCR EIA) testing and conventional cytology or to stand-alone conventional cytology in the first round. In the second round in both arms, HPV testing and cytology were performed on all participants 5 years later. Women with HSIL cytology were immediately referred for colposcopy, and women with ASC-US or LSIL cytology were offered repeat testing after 6 months and 18 months and then referred for colposcopy if they were cytology-positive. In the intervention arm, HPV-positive women with NILM cytology were also offered repeat testing followed by colposcopy if the second HPV test was positive (Bulkmans et al., 2004). Data were initially published on the first two screening rounds, with a 5-year interval, for about half of the cohort (Bulkmans et al., 2007) and then for the entire cohort (Rijkaart et al., 2012a). Further analyses have examined long-term risks (Dijkstra et al., 2016) and additional specific hypotheses on management of different screening results with different combinations of test results over one or two screening rounds (Veldhuijzen et al., 2017; Polman et al., 2019a).

The Randomized Controlled Trial of Human Papillomavirus Testing in Primary Cervical Cancer Screening (SwedeScreen) trial was conducted in five cities in Sweden. A total of 12 527 women aged 32–38 years were enrolled and randomized either to co-testing with HPV

DNA (GP5+/6+ PCR EIA) testing and conventional cytology or to conventional cytology alone (Naucler et al., 2007). Women with ASC-US+ were referred for colposcopy. In the intervention arm, HPV-positive women with NILM cytology received repeat HPV testing after 12 months and were referred for colposcopy if the HPV test result was positive. In the second screening round, all women were screened with conventional cytology. The initial analysis included two screening rounds with an average of 4 years of follow-up per woman. Subsequent analyses have included long-term follow-up data (Elfström et al., 2014; Elfgren et al., 2017).

The A Randomised Trial In Screening To Improve Cytology (ARTISTIC) trial was conducted in Greater Manchester, United Kingdom. A total of 24 510 women aged 20–64 years were enrolled in 2001–2003. Women were randomized 3:1 either to co-testing with HPV DNA (HC2) testing and LBC or to LBC alone. The management of screen-positive women in both arms was similar to that in the POBASCAM trial. The screening protocol for the second round was the same as that for the first round. Data from the first two screening rounds, 3 years apart, were initially reported (Kitchener et al., 2009a, b). Further analyses have reported on the long-term follow-up of this trial (Kitchener et al., 2011).

The Finnish trial was conducted in Finland in 2003–2008 (Leinonen et al., 2012) and enrolled 132 194 women aged 25–65 years. Participants were randomized either to primary screening with HPV DNA (HC2) testing, with conventional cytology triage if HPV-positive (intervention arm) or to conventional cytology alone (control arm). The follow-up period was limited to one screening round with follow-up after 5 years for cumulative detection of CIN, AIS, and invasive cervical cancer. Women in the intervention arm who were HPV-positive and with LSIL or worse (LSIL+) cytology and women in the control arm who were LSIL+ were referred for colposcopy, and women who were HPV-positive and with less than LSIL cytology (intervention arm) or with ASC-US (control arm) were followed up with repeat testing.

The HPV For Cervical Cancer Screening (HPV FOCAL) trial was conducted in Canada in 2008–2016 (Ogilvie et al., 2017, 2018; Coldman et al., 2020). A total of 19 009 women aged 25–65 years attending routine screening were randomized 1:1:1 into one of three groups: primary HPV DNA screening (stand-alone) with LBC triage of HPV-positive women (intervention arm), primary HPV DNA screening (stand-alone) with LBC triage of HPV-positive women and a 2-year safety check (safety arm), and LBC screening with HPV DNA triage of women with an ASC-US result (control arm) and colposcopy for women with LSIL+. In the intervention arm, HPV-negative women were recalled for exit screening with both LBC and HPV testing at 4 years. In the safety arm, HPV-negative women were recalled for exit screening with LBC at 2 years. In the control arm, women with NILM LBC were recalled for screening with LBC at 2 years and then again for exit screening with both LBC and HPV testing at 4 years.

The Hong Kong Special Administrative Region (Hong Kong SAR) trial was conducted at seven clinics in Hong Kong SAR, China, in 2010–2014 (Chan et al., 2020). A total of 15 955 women aged 30–60 years attending routine screening were randomized either to co-testing with HPV testing and LBC (intervention arm) or to LBC with HPV DNA triage of women with an ASC-US+ result (control arm). Women were referred for colposcopy if they were HPV-positive and/or had LSIL+. If the co-testing result was HPV-negative and ASC-US, repeat testing was offered. There were two rounds of screening, with a 3-year interval, and all women were screened with LBC in the second round.

The Compass trial, in Australia, is the first prospective RCT of primary HPV screening compared with cytology to be conducted in a

population with high coverage of HPV vaccination. Women aged 25–64 years were enrolled in 2015–2019 (Canfell et al., 2018). Participants were randomized 1:2 either to 2.5-yearly LBC with HPV triage of low-grade LBC (control arm) or to 5-yearly primary HPV testing (intervention arm). In the intervention arm, women who are positive for HPV16 or HPV18 are directly referred for colposcopy, and women who are positive for other (non-HPV16/18) carcinogenic HPV types undergo secondary randomization 1:1 to either LBC or dual-stain cytology ($p16^{INK4a}$ and Ki-67). In addition, 10% of women in the intervention arm who test negative for HPV will be recalled at 2.5 years for screening with LBC, for safety monitoring purposes. To date, data on the baseline and 12-month follow-up in 4995 women enrolled in 2013–2014 in the Compass pilot trial have been published (Canfell et al., 2017).

The only RCT to evaluate the effect of a single round of screening on cervical cancer incidence and associated mortality was conducted in Osmanabad District in India. This cluster RCT included 131 746 women aged 30–59 years from 52 village clusters randomly assigned to four groups in 2000–2003 (Sankaranarayanan et al., 2009). The groups were randomly assigned to undergo screening with HPV testing (34 126 women), conventional cytology (32 058 women), or VIA (34 074 women) or to receive standard care without screening (31 488 women; control group). Women who had positive results on screening underwent colposcopy and directed biopsies, and those with cervical precancerous lesions or cancer received appropriate treatment. The main results were reported with follow-up until 2007.

Efficacy results from RCTs comparing HPV-based screening with cytology-based screening have been compiled in systematic reviews and meta-analyses (Arbyn et al., 2012; Melnikow et al., 2018). Results per trial are presented in Table 4.22 and in Fig. 4.4. Relative risks and 95% confidence intervals were recalculated by the Working Group. A normal distribution for the logarithm of the estimated relative risk was used to calculate confidence intervals. The NTCC first phase and second phase were pooled, and only NTCC participants aged 35 years and older were included in the analyses. Pooled meta-analytic estimates of the relative risks were calculated by the Working Group assuming a random-effects model and applying restricted maximum-likelihood estimation.

(ii) Detection of CIN2+ and CIN3+

In the eight RCTs comparing primary HPV DNA testing alone or co-testing with HPV DNA testing and cytology (intervention arm) with cytology (control arm), there was consistent evidence that the detection rates of CIN2+ and CIN3+ were higher in the HPV DNA testing arm than in the cytology arm in the first round of screening (Fig. 4.4). In the eight RCTs, the relative risk for the detection of CIN2+ by HPV DNA testing compared with cytology ranged from 1.13 (95% CI, 0.94–1.37) in the ARTISTIC trial (Kitchener et al., 2009b) to 10.95 (95% CI, 1.51–79.34) in the Compass trial (Canfell et al., 2017), and the relative risk for the detection of CIN3+ ranged from 0.97 (95% CI, 0.75–1.25) in the ARTISTIC trial (Kitchener et al., 2009b) to 7.46 (95% CI, 1.02–54.66) in the Compass pilot trial (Canfell et al., 2017). Although the relative risks shown in Fig. 4.4 varied considerably across studies, seven of the eight RCTs reported a relative risk for the detection of CIN2+ with a lower bound of the 95% confidence interval between 1 and 2, and five of the eight RCTs reported a relative risk for the detection of CIN3+ with a lower bound of the 95% confidence interval between 1 and 2.

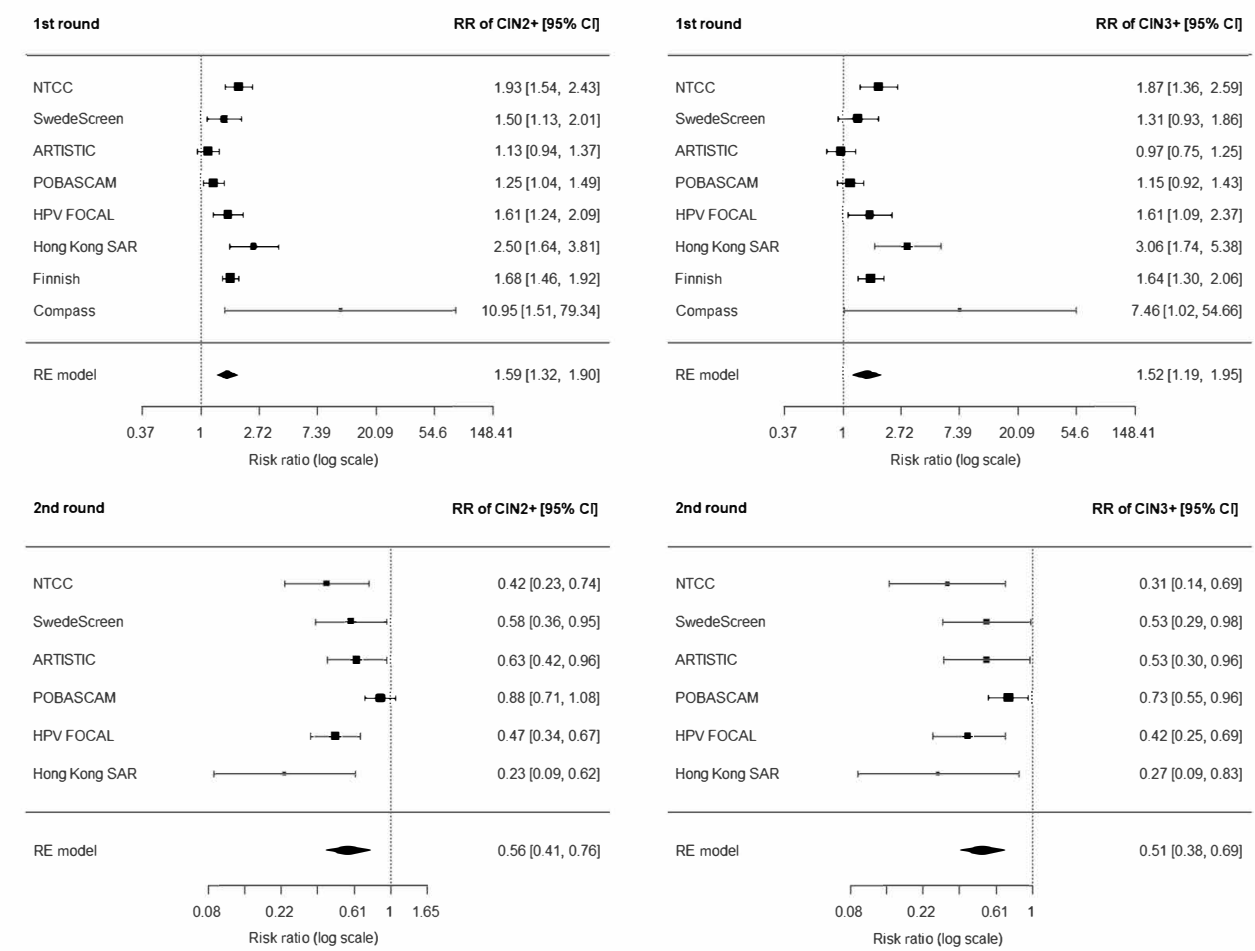
The risk of CIN2+ in the second round of screening was significantly lower in women who were randomized to HPV testing than in those in the cytology arm in the first round of screening (Fig. 4.4). The relative risk of CIN2+ ranged from

Table 4.22 Randomized controlled trials with an HPV-based screening arm (intervention arm) and a cytology arm (control arm)

Trial Country Reference	Age (years)	No. of screening rounds (interval, years)	Screening strategy in round 1: intervention vs control	No. of women in round 1	No. of colposcopy referrals (%)	No. detected (%)		PPV for CIN3+ (%)	No. of women for round 2 calculation	No. detected (%)	
						CIN2+	CIN3+			CIN2+	CIN3+
NTCC Italy Ronco et al. (2006b, 2008, 2010)	35–60	2 (3)	Co-testing (phase 1) or hrHPV (phase 2)	34 430	2768 (8.0%)	213 (0.6%)	105 (0.3%)	3.8	33 733	16 (0.05%)	8 (0.02%)
			Cytology	34 405	928 (2.7%)	110 (0.3%)	56 (0.2%)	6.0	34 202	39 (0.1%)	26 (0.08%)
SwedeScreen Sweden Naucler et al. (2007)	32–38	2 (3)	Co-testing	6257	265 (4.2%)	114 (1.8%)	72 (1.2%)	27.2	6257	25 (0.4%)	16 (0.3%)
			Cytology	6270	150 (2.4%)	76 (1.1%)	55 (0.9%)	36.7	6270	43 (0.7%)	30 (0.5%)
ARTISTIC United Kingdom Kitchener et al. (2009a)	20–64	2 (3)	Co-testing	18 386	1247 (6.8%)	453 (2.5%)	233 (1.3%)	18.7	11 676	65 (0.6%)	29 (0.3%)
			Cytology	6124	320 (5.2%)	133 (2.2%)	80 (1.3%)	25.0	3866	34 (0.9%)	18 (0.5%)
Finnish Finland Leinonen et al. (2012)	25–65	1 (5)	hrHPV	66 410	NR	540 (0.8%)	195 (0.3%)	NR	NR	NR	NR
			Cytology	65 784	NR	319 (0.5%)	118 (0.2%)	NR	NR	NR	NR
POBASCAM Netherlands Rijkaart et al. (2012a)	29–56	2 (5)	Co-testing	19 999	NR	267 (1.3%)	171 (0.9%)	NR	19 579	160 (0.8%)	88 (0.5%)
			Cytology	20 106	NR	215 (1.1%)	150 (0.7%)	NR	19 731	184 (0.9%)	122 (0.6%)
Compass Australia Canfell et al. (2017)	25–64	1 (5)	hrHPV	4000	154 (3.8%)	44 (1.1%)	30 (0.8%)	19.5	NR	NR	NR
			Cytology	995	27 (2.7%)	1 (0.1%)	1 (0.1%)	3.7	NR	NR	NR
HPV FOCAL Canada Ogilvie et al. (2018)	25–65	2 (4)	hrHPV	9540	544 (5.7%)	147 (1.5%)	67 (0.7%)	12.3	9540	48 (0.5%)	22 (0.2%)
			Cytology	9408	290 (3.1%)	90 (9.6%)	41 (0.4%)	14.1	9408	100 (1.1%)	52 (0.6%)
Hong Kong Special Administrative Region trial China Chan et al. (2020)	30–60	2 (3)	Co-testing	7931	738 (9.3%)	75 (1.0%)	49 (0.6%)	6.6	6018	5 (0.08%)	4 (0.07%)
			Cytology	7927	157 (2.0%)	30 (0.4%)	16 (0.2%)	10.2	6203	22 (0.4%)	15 (0.2%)

ARTISTIC, A Randomised Trial In Screening To Improve Cytology; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; HPV FOCAL, HPV For Cervical Cancer Screening; NR, not reported; NTCC, New Technologies for Cervical Cancer Screening; POBASCAM, Population Based Screening Study Amsterdam; PPV, positive predictive value; RCT, randomized controlled trial; SwedeScreen, Randomized Controlled Trial of Human Papillomavirus Testing in Primary Cervical Cancer Screening; yr, year or years.

Fig. 4.4 Randomized controlled trials comparing HPV-based screening versus cytology screening: relative risk of CIN2+ and CIN3+ in the first and second screening rounds



Risk ratio (RR) of CIN2+ (left panel) or CIN3+ (right panel) at first (top) and second (bottom) cervical screening rounds comparing HPV testing with cytology in eight clinical trials.
ARTISTIC, A Randomised Trial In Screening To Improve Cytology; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; HPV FOCAL, HPV For Cervical Cancer Screening; NTCC, New Technologies for Cervical Cancer Screening; POBASCAM, Population Based Screening Study Amsterdam; RE model, random-effects model; SAR, Special Administrative Region; SwedeScreen, Randomized Controlled Trial of Human Papillomavirus Testing in Primary Cervical Cancer Screening.
The pooled estimates were computed by the Working Group based on the data presented in Table 4.22, using the restricted maximum-likelihood estimator method of the metafor library in R for random/mixed-effects models. Source: see Table 4.22 for references.

0.23 (95% CI, 0.09–0.62) in the Hong Kong SAR trial (Chan et al., 2020) to 0.88 (95% CI, 0.71–1.08) in the POBASCAM trial (Rijkaart et al., 2012a), and the relative risk of CIN3+ ranged from 0.27 (95% CI, 0.09–0.83) in the Hong Kong SAR trial (Chan et al., 2020) to 0.73 (95% CI, 0.55–0.96) in the POBASCAM trial (Rijkaart et al., 2012a).

The ARTISTIC, POBASCAM, and SwedeScreen trials also reported the cumulative number of CIN2+ and CIN3+ cases detected in the first and second rounds and during extended follow-up beyond the second round, stratified by the HPV DNA testing and/or cytology result at baseline (Kitchener et al., 2011; Elfström et al., 2014; Dijkstra et al., 2016). In the ARTISTIC trial, the cumulative CIN3+ risk in women with a negative HPV test was 0.13% after two rounds of screening (with an interval of 3 years) and 0.28% after three rounds of screening, whereas the cumulative CIN3+ risk in women with normal cytology was 0.31% after two rounds and 0.63% after three rounds. In the POBASCAM and SwedeScreen trials, separate CIN3+ risks were calculated for the intervention arm and the control arm. In the POBASCAM trial, the cumulative CIN3+ risk in women from the intervention arm with a negative HPV test was 0.31% (95% CI, 0.24–0.41%) after two rounds of screening (with an interval of 5 years) and 0.56% (95% CI, 0.45–0.70%) after three rounds of screening, whereas the cumulative CIN3+ risk in women from the control group with normal cytology was 0.69% (95% CI, 0.58–0.82%) after two rounds and 1.20% (95% CI, 1.01–1.37%) after three rounds (Dijkstra et al., 2016). In the SwedeScreen trial, follow-up data were collected up to 13 years after enrolment and reported for specific time points. The cumulative CIN3+ risk in women from the intervention group with a negative HPV test was 0.04% after 3 years, 0.15% after 5 years, and 0.44% after 10 years, whereas the cumulative CIN3+ risk in women from the control group with normal cytology was 0.20% after 3 years, 0.51% after 5 years, and 0.97% after 10 years (Elfström et al., 2014). The relative cumulative risk of CIN3+ in HPV-negative women compared with women with normal cytology ranged from 0.42 to 0.57 across trials and time points.

[The studies showed considerable variation in HPV and cytology testing technology, age ranges, and management in the HPV DNA testing intervention arms. Five of the eight RCTs evaluated co-testing with HPV testing and cytology compared with cytology alone. The trials also differed in their methods of disease ascertainment at exit testing. For example, in the NTCC and SwedeScreen trials the second round of screening was conducted with cytology, whereas in the POBASCAM and HPV FOCAL trials the second round of screening was conducted with co-testing with HPV testing and cytology, and in the ARTISTIC trial the screening protocols were the same in the first and second rounds. Furthermore, the definition of the second screening round varied across studies. In some trials (e.g. the POBASCAM and HPV FOCAL trials), the start of the second round was based only on time since enrolment, whereas some other trials also used criteria for the start of the second round that depended on the screening results in the first round. Despite design differences, most trials showed an increase in CIN3+ in the first round, and all trials with two screening rounds showed a decrease in CIN3+ in the second round.]

(iii) Efficacy of screening for prevention of cervical cancer and associated death

In the Osmanabad District trial (Sankaranarayanan et al., 2009), different screening strategies (HPV testing, conventional cytology, and VIA) were compared with standard care, but risk ratios for the comparison of HPV testing with cytology can be calculated from the tabulated number of cases and the person-years at risk. The risk ratios for the detection of advanced cancer (International Federation of Gynecology

and Obstetrics [FIGO] stage II or higher) and for cervical cancer mortality in the HPV testing group compared with the cytology group were 0.63 (95% CI, 0.41–0.96) and 0.59 (95% CI, 0.37–0.92), respectively. No reduction in all-cause mortality was observed for any screening intervention group compared with the standard-care control group.

[It is important to bear two issues in mind when interpreting the findings. First, the trial represented the findings of one round of screening in a previously unscreened population. Therefore, risk ratios for cervical cancer mortality are different from those in situations where women are repeatedly screened during their lifetime. Second, although active steps were taken to ascertain vital status and cause of death in the population, it is possible that in this setting there were some limitations in the processes of cancer registration and death ascertainment.]

A pooled analysis of four RCTs conducted in Europe compared the efficacies of HPV DNA testing and cervical cytology for the prevention of invasive cervical cancer (Ronco et al., 2014). This analysis was critical, because it examined an invasive cervical cancer end-point for the first time in a high-income country setting. The pooled analysis included 176 464 women aged 20–64 years who were randomly assigned to HPV-based screening (intervention arm) or cytology-based screening (control arm) in Italy (NTCC), the Netherlands (POBASCAM), Sweden (SwedeScreen), and the United Kingdom (ARTISTIC). Women were followed up for a median of 6.5 years, and during that time 107 invasive cervical carcinomas were detected. Cumulative detection of invasive cervical cancer was lower in the HPV testing arm than in the cytology arm during the study period (rate ratio, 0.60; 95% CI, 0.40–0.89), and no heterogeneity was detected between studies ($P = 0.52$). Detection of invasive cervical carcinoma was similar between screening methods during the first 2.5 years of follow-up (rate ratio, 0.79; 95% CI, 0.46–1.36) but was significantly lower in the HPV arm thereafter (rate ratio, 0.45; 95% CI, 0.25–0.81). In women with a negative screening test at entry (HPV-negative in the intervention arm and cytology-negative in the control arm), the rate ratio was 0.30 (95% CI, 0.15–0.60). The cumulative incidence of invasive cervical carcinoma in women with negative entry tests was 4.6 (95% CI, 1.1–12.1) per 100 000 women at 3.5 years and 8.7 (95% CI, 3.3–18.6) per 100 000 women at 5.5 years in the HPV testing arm and 15.4 (95% CI, 7.9–27.0) per 100 000 women at 3.5 years and 36.0 (95% CI, 23.2–53.5) per 100 000 women at 5.5 years in the cytology arm. The pooled rate ratio was lower for adenocarcinoma (0.31; 95% CI, 0.14–0.94) than for SCC (0.78; 95% CI, 0.49–1.25). The lowest rate ratios were observed in women aged 30–34 years (0.36; 95% CI, 0.14–0.94).

[The authors found no heterogeneity in efficacy between studies, which supports the pooling of data and the overall pooled findings. It should be noted that data from these trials are representative of women followed up for at least two rounds of screening, which may be different from long-term, steady-state effects of repeated rounds of screening with a particular screening test and management protocol in a population.]

(iv) Harms

Harms during the first round of screening were measured by the proportion of women referred for colposcopy after a positive screening test and by the PPV for CIN3+ (the proportion of CIN3+ detected in women referred for colposcopy). The number of colposcopy referrals includes women who were referred at baseline or after repeat testing within the same screening round. The proportion of colposcopy referrals was generally higher for HPV-based screening than for cytology-based screening (Table 4.22). The biggest differences in colposcopy referrals between the study arms were found in the NTCC trial (8.0% vs 2.7%) and the Hong Kong SAR trial (9.3% vs 2.0%), in which HPV-positive

women were not offered triage testing but were immediately referred for colposcopy. The PPV for CIN3+ was similar in the two study arms or higher in the cytology arm in all studies, with the exception of the Compass trial, in which the PPV was higher in the HPV-based testing arm (19.5%) than in the cytology arm (3.7%).

[The number of women with a positive screening test result and the number of colposcopies should be interpreted in relation to the number of CIN3+ detected. If the number of CIN3+ is proportional to the number of colposcopy referrals, then the harms per detected CIN3+ remain unchanged.]

A more complete picture of the harms of screening is obtained from the number of diagnostic procedures when measured over multiple rounds of screening. In the HPV FOCAL trial, the cumulative colposcopy referral rates were similar in the two study arms over two rounds of screening, and in the Hong Kong SAR trial, in which HPV-positive women were immediately referred for colposcopy, the cumulative colposcopy referral rate was higher in the HPV testing arm than in the cytology arm (relative colposcopy referral rate, 2.83; 95% CI, 2.47–3.24). Similar results on cumulative biopsy rates were observed in four RCTs conducted in Europe (Ronco et al., 2014). In the ARTISTIC, POBASCAM, and SwedeScreen trials, the cumulative biopsy rate over two rounds of screening was similar in the two study arms, whereas in the NTCC trial, in which HPV-positive women were immediately referred for colposcopy, the biopsy rate was higher in the HPV testing arm than in the cytology arm (relative biopsy rate, 2.24; 95% CI, 2.09–2.39).

An indication of overtreatment of cervical lesions can be obtained by comparing the cumulative detection of CIN2+ between the HPV testing arm and the cytology arm over two screening rounds. The relative risks of CIN2+ can be computed from the numbers in Table 4.22. The relative risk of CIN2+ over two screening rounds (as computed by the Working Group) was 1.01 (95% CI, 0.83–1.23) in the HPV FOCAL trial, 1.03 (95% CI, 0.87–1.23) in the ARTISTIC trial, 1.08 (95% CI, 0.94–1.24) in the POBASCAM trial, and 1.17 (95% CI, 0.92–1.49) in the SwedeScreen trial, suggesting that replacing cytology-based screening with HPV-based screening will lead to only a small increase in overtreatment. In the NTCC trial and the Hong Kong SAR trial, the estimated relative risks of CIN2+ over two screening rounds were 1.54 (95% CI, 1.25–1.89) and 1.54 (95% CI, 1.09–2.18), respectively, suggesting a moderate increase in overtreatment.

[A difference in the detection of CIN2+ between study arms over two screening rounds needs to be interpreted with care. It may indicate that the magnitude of overtreatment of CIN2+ differs between study arms, but it may also simply point at a difference in lead-time gain that is longer than the interval between two consecutive screens. In the POBASCAM and HPV FOCAL trials, in which women in both study arms received co-testing in the second screening round, so that differences in lead-time gain have become minimal after the second round , there was no marked difference in cumulative detection of CIN2+ between study arms over two screening rounds.]

(d) Population-based cohorts

(i) Description

Studies in Argentina (Arrossi et al., 2019), Denmark (Thomsen et al., 2020), Finland (Veijalainen et al., 2019), Italy (Pasquale et al., 2015; Maggino et al., 2016; Passamonti et al., 2017; Zorzi et al., 2017), the Netherlands (Aitken et al., 2019), Sweden (Lamin et al., 2017), and the United Kingdom (Rebolj et al., 2019) have reported on the impact of primary HPV DNA screening in national, regional, or pilot screening programmes on precancer and cancer. In all cohort studies, HPV DNA-positive women were triaged with cytology to improve the balance between benefits

and harms. There was considerable variation with respect to the follow-up of HPV-positive women with NILM cytology, who were followed up with cytology in the Netherlands, with HPV testing in Argentina, Finland, and Italy, and with combined HPV testing and cytology in Denmark and the United Kingdom, and were re-invited at the next screening round in Sweden. The studies in Argentina, Finland, Italy, and the Netherlands compared primary HPV-based screening programmes with the cytology-based screening programmes that were offered before the implementation of HPV screening. The study in the United Kingdom compared a pilot HPV-based screening implementation cohort with a cytology-based programme running in the same period and region, and the studies in Denmark and Sweden conducted a randomized health services trial with a primary HPV-based screening arm and a cytology-based screening arm.

Co-testing with HPV testing and cytology has been implemented as a screening option in the USA. In 2003, Kaiser Permanente Northern California (KPNC), a large health maintenance organization, adopted screening based on co-testing, with a 3-year interval after a double-negative screening result. The KPNC cohort comprises about 1 million women aged 30–64 years who have received up to four rounds of co-testing (Castle et al., 2019). Co-testing has also been implemented as a pilot programme in the Wolfsburg region in Germany: the Wolfsburg Pilot Project for Better Prevention of Cervical Cancer with Primary HPV Screening (WOLPHSCREEN). By 2016, the WOLPHSCREEN programme had enrolled 26 624 women aged 30–70 years (Horn et al., 2019). The WOLPHSCREEN programme has a 5-year screening interval after a double-negative screening result. In 2019, women had completed up to three screening rounds. Co-testing cohorts do not have a control group, but comparisons between HPV testing and cytology screening can be made on the basis of the co-testing results. These comparisons are particularly suitable for determining screening intervals (Katki et al., 2011). Further study features of the primary HPV testing and co-testing cohorts, such as study size, age range, and follow-up protocol for HPV DNA-positive women, are given in Table 4.23.

Several other studies have been conducted with one round of co-testing followed by cytology screening in subsequent rounds. These include a pooled analysis of seven studies in European countries (Dillner et al., 2008), including 24 295 women followed up until 6 years after HPV testing who had at least one cervical cytology or histopathology examination during follow-up. Four other studies with a single round of co-testing are available: (i) the HPV in Addition to Routine Testing (HART) study, including 8735 women aged 30–60 years at five clinical centres in the United Kingdom, with a median follow-up of 6 years (Mesher et al., 2010); (ii) the Canadian Cervical Cancer Screening Trial (CCCaST) study, including 4400 women aged 30–69 years in Montreal, with a median follow-up of 1.5 years, and 5754 women aged 30–69 years in St. John's, with a maximum follow-up of 10 years (Isidean et al., 2016); (iii) the Vrije Universiteit Medical Centre-Saltro Laboratory Population-Based Cervical Screening (VUSA-Screen) study, including 25 871 women aged 29–61 years in Utrecht in the Netherlands, with a maximum follow-up of 3 years (Rijkaart et al., 2012b); and (iv) the Addressing the Need for Advanced HPV Diagnostics (ATHENA) study, including 41 955 women aged 25 years and older at 61 clinical centres in the USA, with a follow-up of 3 years (Wright et al., 2015).

(ii) Detection of CIN2+ and CIN3+

The results of the primary HPV screening cohorts with cytology triage for HPV DNA-positive women were consistent with those of the RCTs, because the detection rates of CIN2+ and CIN3+ were always at least as high

Table 4.23 Population-based cohorts: comparison of screening with HPV DNA testing alone or with co-testing versus cytology

Country Reference	Type of study	No. of screened subjects Age (years)	Colposcopy referral recommendation	HPV DNA+/ co-test+ (%)	HPV versus cytology, RR[a] (95% CI)				
					Test-positive	Colposcopy referral	CIN2+	CIN3+	PPV for CIN3+
Argentina Arrossi et al. (2019)	Primary HPV with cytology triage, regional programme (Jujuy)	49 565 30–60	ASC-US, HPV+ at 18 mo	13.6	3.42 (3.22–3.64)	2.69[b] (2.42–2.99)	1.76 (1.52–2.03)	1.90 (1.61–2.24)	1.13 (1.00–1.29)
Denmark Thomsen et al. (2020)	Primary HPV with cytology triage, randomized pilot implementation	11 339 30–59	ASC-US, HPV16/18+, HPV+ or ASC-US at 12 mo	8.8	3.84 (3.42–4.30)	1.81[b] (1.58–2.07)	1.51[b] (1.21–1.89)	1.40 [b] (1.07–1.82)	0.77 (0.62–0.97)
Finland Veijalainen et al. (2019)	Primary HPV with cytology triage, regional programme (Tampere)	17 770 35–60	LSIL, HPV+ or LSIL at 12 mo	8.2	1.10 (1.02–1.19)	1.98 (1.75–2.24)	2.45 (1.76–3.41)	2.70 (1.75–4.17)	1.36 (0.90–2.06)
Germany Luyten et al. (2014)	WOLPHSCREEN cohort. Co-testing, regional pilot programme (Wolfsburg)	19 795 30–70	HPV+ and ASC-US, ASC-US at 6 mo, HPV+ at 12 mo	7.5	2.76 (2.51–3.04)	3.22 (2.87–3.60)	2.50 (2.17–2.87)	2.25 (1.90–2.66)	0.70 (0.59–0.83)
Italy Pasquale et al. (2015)	Primary HPV with cytology triage, regional programme (Valcamonica)	18 728 25–64	ASC-US, HPV+ at 12 mo	8.7	2.33 (2.14–2.54)	1.71 (1.56–1.88)	1.59 (1.23–2.07)	NR	NR
Italy Maggino et al. (2016)	Primary HPV with cytology triage, regional programme (Venice)	89 217 25–64	ASC-US, HPV+ at 12 mo	6.8	2.35 (2.25–2.46)	1.78 (1.70–1.87)	2.23 (1.87–2.65)	NR	NR
Italy Passamonti et al. (2017)	Primary HPV with cytology triage, regional programme (Perugia)	6272 25–64	ASC-US, HPV+ at 12 mo	6.3	4.19 (3.57–4.92)	4.00 (3.29–4.87)	2.65 (1.85–3.78)	NR	NR

Table 4.23 (continued)

Country Reference	Type of study	No. of screened subjects Age (years)	Colposcopy referral recommendation	HPV DNA+/ co-test+ (%)	HPV versus cytology, RR[a] (95% CI)				
					Test-positive	Colposcopy referral	CIN2+	CIN3+	PPV for CIN3+
Italy Zorzi et al. (2017)	Primary HPV with cytology triage, regional programme (Padua)	48 763 25–64	ASC-US, HPV+ at 12 mo	6.4	NR	NR	1.2 (0.9–1.7)	NR	NR
Netherlands Aitken et al. (2019)	Primary HPV with cytology triage, national programme	454 573 29–61	ASC-US, ASC-US at 6 mo	9.1[c]	1.89 (1.86–1.92)	1.97 (1.92–2.02)	1.34 (1.29–1.39)	1.28 (1.23–1.35)	0.65 (0.63–0.68)
Sweden Lamin et al. (2017)	Primary HPV with cytology triage, randomized pilot implementation (Stockholm)	7325 56–60	ASC-US	5.5	2.69 (2.24–3.23)	1.18 (0.81–1.71)	1.07 (0.56–2.04)	1.02 (0.47–2.19)	0.86 (0.44–1.69)
United Kingdom Rebolj et al. (2019)	Primary HPV with cytology triage, non-randomized pilot implementation	183 970 24–64	ASC-US, HPV+ and ASC-US at 12 mo, HPV+ at 24 mo	12.7	3.31 (3.25–3.38)	1.85 (1.80–1.89)	1.46 (1.40–1.52)	1.41 (1.34–1.48)	0.76 (0.73–0.80)
USA Castle et al. (2019)	KPNC cohort. Co-testing, regional cohort (Northern California)	990 013 30–64	LSIL, HPV+ and ASC-US, HPV+ or ASC-US at 12 mo	8.0	1.30 (1.29–1.32)	NR	NR	1.36 (1.33–1.39)	NR

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; HPV, human papillomavirus; KPNC, Kaiser Permanente Northern California; LSIL, low-grade squamous cell intraepithelial lesion; mo, month or months; NR, not reported; PPV, positive predictive value; RR, relative risk; WOLPHSCREEN, Wolfsburg Pilot Project for Better Prevention of Cervical Cancer with Primary HPV Screening.

[a] The relative risks, computed by the Working Group, are based on absolute numbers reported in the original publications. The 95% confidence intervals were calculated using a normal reference distribution for the logarithm of the estimated relative risk.

[b] Baseline only; no repeat testing information used.

[c] Absolute numbers were not available; based on proportions reported in the article.

with HPV screening as with cytology screening (Table 4.23). In studies that reported on both CIN2+ and CIN3+ cases, the relative risks of HPV testing versus cytology were similar for both end-points. The relative risks for the detection of CIN2+ varied from 1.07 (95% CI, 0.56–2.04) in the study in Sweden (restricted to women aged 56–60 years) to 2.65 (95% CI, 1.85–3.78) in the study in Perugia in Italy.

[In the studies in Argentina and Denmark, follow-up data for HPV-positive women with NILM cytology were incomplete. This may have led to an underestimation of the relative detection risk, because women with NILM cytology have a relatively low CIN2+ risk.]

Most countries implemented primary HPV screening with cytology triage in women older than 30 years, but in some regions in Italy and in the United Kingdom, HPV screening was also studied in women aged from 24 or 25 years to 29 years. In the areas of Padua, Valcamonica, and Venice in Italy, the risks of CIN2+ per screened woman were 1.0%, 2.1%, and 1.1%, respectively, in women younger than 30 years and 0.4%, 0.6%, and 0.4%, respectively, in women aged 30 years and older (Pasquale et al., 2015; Maggino et al., 2016; Zorzi et al., 2017). In the pilot implementation cohort in the United Kingdom, the risk of CIN2+ per screened woman was 6.6% in women younger than 30 years and 1.2% in women aged 30 years and older, and the risk of CIN3+ per screened woman was 4.0% in women younger than 30 years and 0.8% in women aged 30 years and older (Rebolj et al., 2019).

This risks of CIN2+ and CIN3+ in subsequent screening rounds were also studied in the cohorts in Italy. In the cohort in Padua (Zorzi et al., 2017), the CIN2+ risk in the second round after 3 years was 0.11% per screened woman and the CIN3+ risk was 0.03%. The relative risk of CIN2+ in the second round versus the first round was 0.24 (95% CI, 0.16–0.37), and the relative risk of CIN3+ was 0.14 (95% CI, 0.06–0.32). In the cohort in Perugia (Passamonti et al., 2017), the risks of CIN2+ and CIN3+ in the second round after 3 years were 0.25% and 0.17%, respectively, and the relative risks of CIN2+ and CIN3+ were 0.25 (95% CI, 0.14–0.42) and 0.39 (95% CI, 0.20–0.79), respectively. In a study of three cohorts in Italy (Del Mistro et al., 2019), the relative risks of CIN2+ and CIN3+ in the second round versus the first round were found to be higher when an HPV infection was reported in the previous round, and also when the positive HPV test result was followed by a negative HPV test result during short-term repeat testing. This finding was also reported for the intervention arm of the POBASCAM trial (Polman et al., 2017).

[The low risks of CIN2+ and CIN3+ in the second primary HPV screening round support the use of intervals of longer than 3 years when the primary HPV test result in the previous round is negative.]

Table 4.23 also shows the results of the cohorts in which co-testing with HPV testing and cytology has been implemented: the WOLPHSCREEN cohort in Germany and the KPNC cohort in the USA. For both studies, substantially higher CIN3+ risks were observed after a positive HPV test result than after abnormal cytology. In addition, in the KPNC cohort, the 5-year CIN3+ risk was 0.11% after a negative HPV test result and 0.25% after an NILM cytology result (Castle et al., 2018). In the WOLPHSCREEN cohort, the 5-year CIN3+ risk was 0.013% after a negative HPV test result and 0.071% after an NILM cytology result (Horn et al., 2019). Cohorts with only one round of co-testing followed by cytology follow-up yielded results that were in line with those from the KPNC and WOLPHSCREEN cohorts. In a pooled study of seven European cohorts (Dillner et al., 2008), the pooled 5-year CIN3+ risk was 0.27% after a negative HPV test result and 0.83% after an NILM cytology result. The VUSA-Screen study reported a 3-year CIN3+ risk of 0.06% after a negative HPV test result and 0.26% after NILM

cytology, and the ATHENA study reported a 3-year CIN3+ risk of 0.3% after a negative HPV test result and 0.8% after NILM cytology. The HART study and the CCCaST study reported risks only for the end-point CIN2+. In the HART study, the 3-year CIN2+ risk was 0.04% after a negative HPV test result and 0.21% after NILM cytology, and the 5-year CIN2+ risk was 0.15% after a negative HPV test result and 0.28% after NILM cytology. In the CCCaST study, the 3-year CIN2+ risk was 0.90% after a negative HPV test result and 1.40% after NILM cytology.

(iii) Detection of cervical cancer

The two largest primary HPV screening cohorts, in the United Kingdom (Rebolj et al., 2019) and the Netherlands (Aitken et al., 2019), reported on cervical cancer detection over one round of screening and compared it with the cancer detection in a historical cytology screening cohort. In the cohort in the United Kingdom, cervical cancer detection over one round of screening was 0.05% for HPV DNA screening and 0.04% for cytology screening, and the adjusted odds ratio for cervical cancer detection was 1.27 (95% CI, 0.99–1.63) (Rebolj et al., 2019). In the cohort in the Netherlands, cervical cancer detection over one round was 0.04% for HPV DNA screening and 0.03% for cytology screening (Aitken et al., 2019).

In the KPNC co-testing cohort, the 5-year cancer risk was 0.5% after a positive HPV DNA test result and 0.5% after abnormal cytology (Castle et al., 2019). In the subgroup of women with a negative HPV test result (Castle et al., 2018), the 5-year cancer risk was 0.009%, which was about 40% lower than the 5-year cancer risk of 0.02% after an NILM cytology result. The cancer risk after a negative HPV test result further decreased after previous rounds of negative HPV testing: the 5-year cancer risk was 0.004% after two rounds of negative HPV DNA testing and 0.002% after three rounds of negative HPV DNA testing. The results from the KPNC cohort were supported by the findings of the WOLPHSCREEN study, in which the risk of cancer in the first co-testing screening round was 0.10%, which further decreased to 0.03% in subsequent rounds (Horn et al., 2019).

[Together, the RCTs, the primary HPV screening cohorts, and the co-testing cohorts demonstrate that a negative HPV test result gives better reassurance against CIN3+ and cancer than does NILM cytology, and supports the use of longer screening intervals.]

(iv) Harms

In the primary HPV screening cohorts, both the proportion of screen-positive women and the proportion of colposcopy referrals were higher than in cytology screening cohorts (Table 4.23). However, the proportions varied widely across studies. The relative proportion of screen-positive women varied from 1.10 (95% CI, 1.02–1.19) in the study in Finland to 3.84 (95% CI, 3.42–4.30) in the study in Denmark, and the relative proportion of colposcopy referrals varied from 1.18 (95% CI, 0.81–1.71) in the study in Sweden to 4.00 (95% CI, 3.29–4.87) in the study in Perugia in Italy. The proportion of CIN3+ per colposcopy referral (PPV for CIN3+) was below 1 in most settings (up to 35% lower in the Netherlands) but was higher in the studies in Argentina (RR, 1.13; 95% CI, 1.00–1.29) and in Finland (RR, 1.36; 95% CI, 0.90–2.06). In Italy, the studies in Perugia (Passamonti et al., 2017) and in Padua (Zorzi et al., 2017) also reported on the colposcopy referrals in the second HPV-based screening round. The proportion of colposcopy referrals per screened woman in the second round decreased by 10% (95% CI, −6% to 25%) in the Perugia cohort and by 51% (95% CI, 46–55%) in the Padua cohort compared with the first HPV-based screening round. The proportion of CIN3+ per colposcopy referral decreased by 58% (95% CI, 17–78%) in the Perugia cohort and by 71% (95% CI, 35–87%) in the Padua cohort.

[It must be recognized that the follow-up of HPV-positive women with NILM cytology was incomplete in the studies in Argentina and Denmark, and that in Sweden, HPV-positive women with NILM cytology did not receive short-term follow-up testing. This may influence the proportion of colposcopy referrals, which was lowest in Sweden. The high PPV for CIN3+ in the study in Finland is a direct consequence of the high relative detection rate of CIN3+ per screened woman in this study, which was the highest among the studies that reported on CIN3+ cases.]

Consistent with results from the primary HPV screening cohorts, the proportion of screen-positive women was higher for HPV testing than for cytology in the two co-testing cohorts (KPNC and WOLPHSCREEN). The WOLPHSCREEN cohort also reported that the number of colposcopy referrals in HPV-positive women was 3.22 (95% CI, 2.87–3.60) times that in women with abnormal cytology; the corresponding relative PPV for CIN3+ after colposcopy referral was 0.70 (95% CI, 0.59–0.83).

[Both triage testing of HPV-positive women and suitable follow-up management of HPV-positive women with NILM cytology results are important to achieve a good balance between screening benefits and harms. Nonetheless, the results from population-based cohorts indicated that an increase in the number of colposcopy referrals can be expected in the first round of HPV-based screening.]

4.4.3 Comparison of HPV DNA testing versus VIA

(a) Introduction

No review was available that directly compared the impact of HPV DNA testing and VIA on cervical cancer incidence, mortality, and detection.

Evidence about diagnostic accuracy was extracted from eight reviews and meta-analyses or pooled analyses across a wide range of geographical regions. Data were drawn from observational studies, and mostly cross-sectional studies; this may limit the strength of the evidence. In addition, the original studies included in the reviews and analyses had not necessarily compared HPV DNA testing and VIA directly. Thus, the pooled results may potentially be affected by multiple factors, including but not limited to (i) non-comparability of control groups, (ii) different screening participation rates across studies, and (iii) heterogeneity in quality assurance and monitoring methods. Moreover, the performance of VIA, which is a technique that is highly subjective and heavily dependent on the training and experience of providers, varied widely across different populations and research settings (see Sections 4.2.1–4.2.3). In addition, in many studies in which VIA was evaluated, colposcopy plus directed biopsy used as the reference were generally applied to women with a positive screening test result only, potentially leading to verification bias. Furthermore, colposcopy could miss up to 40% of prevalent precancers and is closely correlated with visual screening approaches (see Section 4.2.2); such potential outcome misclassification with VIA may greatly affect the estimates of the test accuracy. Given the above-mentioned limitations, in comparisons of HPV DNA testing with VIA, the results for accuracy parameters must be interpreted with caution.

The detection rate of cervical neoplasia and cancer was assessed mainly by two RCTs, a pooled analysis of two cohort studies, and three cross-sectional studies, one of which was applied in a real-world setting in China.

The incidence of and mortality from cervical cancer were assessed by an RCT in Osmanabad District in India, which was the only study available.

Table 4.24 Accuracy of HPV DNA testing versus visual inspection with acetic acid (VIA)

Reference Study population	Screening exposure Age of included subjects (years) End-point	Test positivity rates (%) (95% CI)	Sensitivity estimate, % (95% CI)		Specificity estimate, % (95% CI)		Relative sensitivity (95% CI) Relative specificity (95% CI)	Comments
			HPV	VIA	HPV	VIA		
Arbyn et al. (2008) Pooled analysis of > 58 000 women aged 25–64 yr recruited from 11 cross-sectional studies in urban settings in India and French-speaking countries in Africa in 1999–2003	HPV DNA test, VIA, VILI, VIAM, cytology (see comments) 25–64 CIN2+, CIN3+, cancer	VIA: 16.7; range, 6.0–27.4	CIN2+: 61.9 (56.2–67.7); range, 48.4–67.7 CIN3+: 68.4 (61.5–75.4); range, 62.3–73.5 Cancer: 72.1 (60.3–83.8); range, 61.5–85.7	CIN2+: 79.2 (73.3–85.0); range, 65.0–91.1 CIN3+: 82.9 (77.1–88.7); range, 58.3–94.6 Cancer: 88.7 (83.1–94.3); range, 66.7–100.0	CIN2+: 93.6 (92.4–94.8); range, 91.6–94.6 CIN3+: 93.4 (92.2–94.6); range, 91.4–94.4 Cancer: 93.0 (91.8–94.2); range, 91.4–94.0	CIN2+: 84.7 (80.7–88.0); range, 74.2–94.5 CIN3+: 84.2 (80.0–88.3); range, 73.8–94.3 Cancer: 83.6 (79.3–88.0); range, 73.1–94.1	HPV vs VIA: CIN2+: 0.883 (0.775–1.007) CIN3+: 0.956 (0.781–1.169) HPV vs VIA: CIN2+: 1.074 (1.051–1.097) CIN3+: 1.075 (1.051–1.099)	Evidence from observational studies. Not every study included had assessed the HPV DNA test and VIA concurrently. HPV DNA test (HC2) was applied in 4 studies in India, and VIA was used in all 11 studies in both Africa and India
Zhao et al. (2010) Pooled analysis of individual patient data in 28 848 women from 17 population-based, cross-sectional cervical cancer screening studies in both urban and rural areas in 9 provinces in China in 1999–2008. The eligible women were sexually active, were not pregnant, had an intact uterus, and had no history of CIN or cervical cancer	HPV DNA test, VIA, cytology 17–59 CIN2+, CIN3+	HPV: 16.3 (4691 of 28 848 women) VIA: 10.8 (3122 of 28 815 women)	Uncorrected: CIN2+: 96.3 (94.9–97.4) CIN3+: 97.5 (95.7–98.7) Corrected: CIN2+: 95.1 (93.6–96.3) CIN3+: 97.6 (95.9–98.6)	CIN2+: 48.0 (42.1–53.9); range, 12.5–70.2 CIN3+: 54.6 (48.0–61.2); range, 14.3–85.7	Uncorrected: CIN2+: 86.4 (83.8–89.0) CIN3+: 85.1 (82.3–87.9) Corrected: CIN2+: 85.4 (85.0–85.8) CIN3+: 84.1 (83.7–84.5)	CIN2+: 90.4 (87.3–93.5); range, 70.0–98.2 CIN3+: 89.9 (86.8–93.0); range, 69.9–97.5	NR NR	Evidence from observational studies. Women included in the pooled analysis all concurrently received HPV DNA test, LBC, and VIA

Table 4.24 (continued)

Reference Study population	Screening exposure Age of included subjects (years) End-point	Test positivity rates (%) (95% CI)	Sensitivity estimate, % (95% CI)		Specificity estimate, % (95% CI)		Relative sensitivity (95% CI) Relative specificity (95% CI)	Comments
			HPV	VIA	HPV	VIA		
Chen et al. (2012) 101 299 apparently healthy women from 22 cross-sectional studies (99 972 women tested by VIA, 23 628 women tested by HPV DNA test). 6 common cervical screening strategies including VIA and HPV DNA test were assessed	HPV DNA test, VIA, VIAM, VILI, cytology (see comments) 16–70 CIN2+	NR	74 (69–78)	77 (75–78)	92 (92–93)	87 (87–88)	NR NR	Studies included in the review underwent quality assessment with QUADAS and STARD quality assessment criteria. Evidence from observational studies. Not every study included had assessed the HPV DNA test and VIA concurrently. Three types of HPV DNA test were involved (HC2, PCR, and *care*HPV), but only the HC2 assay with samples collected by health professionals was used to estimate the accuracy of HPV testing in this meta-analysis

Table 4.24 (continued)

Reference Study population	Screening exposure Age of included subjects (years) End-point	Test positivity rates (%) (95% CI)	Sensitivity estimate, % (95% CI)		Specificity estimate, % (95% CI)		Relative sensitivity (95% CI) Relative specificity (95% CI)	Comments
			HPV	VIA	HPV	VIA		
Fokom-Domgue et al. (2015) 8 studies in which the reference standard (colposcopy and colposcopy-directed biopsy) was performed in all women of the study population from sub-Saharan Africa were included. The study population was not at particular risk of cervical cancer (studies focusing on HIV-positive women or on women presenting with gynaecological symptoms were excluded). In total, 47 361 women were screened with VIA and 3950 women were screened with HPV DNA test	HPV DNA test, VIA, VILI 15–83 CIN2+	HPV: 25.8 (17.4–35.3); range, 12.5–42.8 VIA: 16.8 (11.0–23.6); range, 3.1–39.9	88.3 (73.1–95.5); range, 80.2–96.2	82.4 (76.3–87.3); range, 65.0–94.4	73.9 (50.7–88.7); range, 61.2–88.9	87.4 (77.1–93.4); range, 64.1–98.2	VIA vs HPV: 0.94 (0.82–1.16) VIA vs HPV: 1.17 (0.95–1.69)	Studies included were assessed as of moderate quality, based on the QUADAS-2 criteria. Evidence from observational studies. Not every study included had assessed the HPV DNA test and VIA concurrently. Test accuracy was assessed only among the studies in which the reference test (colposcopy and colposcopy-directed biopsy) was performed in all women (10 studies for VIA, 3 studies for HPV), which may avoid verification bias

Table 4.24 (continued)

Reference Study population	Screening exposure Age of included subjects (years) End-point	Test positivity rates (%) (95% CI)	Sensitivity estimate, % (95% CI)		Specificity estimate, % (95% CI)		Relative sensitivity (95% CI) Relative specificity (95% CI)	Comments
			HPV	VIA	HPV	VIA		
Bobdey et al. (2015) 16 studies conducted in India in 1990–2013 were included. Pooled data of 89 461 women in the VIA arm from 14 studies and 23 244 women in the HPV test arm from 8 studies were analysed	HPV DNA test, VIA, VIAM, VILI, cytology NA NR	NR	75.04; range, 45.70–97.10	68.76; range, 31.60–100.00	91.66; range, 84.20–94.60	84.02; range, 53.30–91.23	NR NR	No quality assessment criteria were applied in the 2 reviews. The age range of included participants and disease end-points of assessment, and the 95% CI of the pooled results on accuracy were not reported. Evidence from observational studies. Not every study included had assessed the HPV DNA test and VIA concurrently. Some included studies were conducted in the health clinics including gynaecologically symptomatic women. Thus, the pooled results of accuracy in the reviews consisted of both asymptomatic and symptomatic participants, which may limit the generalizability to healthy women
Bobdey et al. (2016) 11 studies conducted in India in 1990–2015 were included. Pooled number of women in the VIA arm was 57 225 and in the HPV DNA test arm was 25 575	HPV DNA test, VIA, VIAM, VILI, cytology NA NR	NR	77.81	67.65	91.54	84.32	NR NR	

Table 4.24 (continued)

Reference Study population	Screening exposure Age of included subjects (years) End-point	Test positivity rates (%) (95% CI)	Sensitivity estimate, % (95% CI)		Specificity estimate, % (95% CI)		Relative sensitivity (95% CI) Relative specificity (95% CI)	Comments
			HPV	VIA	HPV	VIA		
Mustafa et al. (2016) 5 cross-sectional studies with a total of 8921 non-pregnant women not previously diagnosed with cervical neoplasia were included	HPV DNA test, VIA, cytology ≥ 18 CIN2/3	HPV: 17.6 VIA: 14.1	95 (84–98); range, 64–97	69 (54–81); range, 41–87	84 (72–91); range, 56–93	87 (79–92); range, 76–95	NR NR	All the included studies underwent quality assessment with QUADAS criteria. Evidence from observational studies. Women included in the studies had all concurrently received HPV DNA test and VIA
Holt et al. (2017) Data of 2757 postmenopausal women were extracted from the 17 population-based studies in Zhao et al. (2010) for further analysis	HPV DNA test, VIA, cytology 17–59 CIN2+, CIN3+	HPV: 17.2 (15.9–18.7) VIA: 6.2 (5.3–7.1)	CIN2+: 82/84, 97.6 (92.4–99.6) CIN3+: 47/48, 97.9 (90.2–99.9)	CIN2+: 26/84, 31.0 (21.8–41.4) CIN3+: 20/48, 41.7 (28.4–55.9)	CIN2+: 2280/2673, 85.3 (83.9–86.6) CIN3+: 2281/2709, 84.2 (82.8–85.5)	CIN2+: 2529/2673, 94.6 (93.7–95.4) CIN3+: 2559/2709, 94.5 (93.6–95.3)	NR NR	This is a further stratification analysis after the pooled analysis of 17 cross-sectional studies described in Zhao et al. (2010)

CI, confidence interval, CIN, cervical intraepithelial neoplasia; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HC2, Hybrid Capture 2; HPV, human papillomavirus; LBC, liquid-based cytology; mo, month or months; NR, not reported; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; STARD, Standards for Reporting of Diagnostic Accuracy Studies; VIA, visual inspection with acetic acid; VIAM, visual inspection with acetic acid using low-level magnification; VILI, visual inspection with Lugol's iodine.

(b) Accuracy of HPV DNA testing versus VIA

Studies comparing the accuracy of HPV DNA testing versus VIA are presented in Table 4.24. Most of the reviews reported a higher pooled sensitivity for HPV DNA testing compared with VIA, and the clinical performance of VIA varied greatly across different geographical areas and studies, which highlighted the difficulties in achieving reliable performance of VIA (Arbyn et al., 2008; Zhao et al., 2010; Chen et al., 2012; Bobdey et al., 2015, 2016; Fokom-Domgue et al., 2015; Mustafa et al., 2016). The sensitivity of HPV DNA testing for detection of CIN2+ varied from 61.9% with HC2 test data pooled from studies in India (Arbyn et al., 2008) to 96.3% in the pooled analysis in China (Zhao et al., 2010); the sensitivity of VIA for detection of CIN2+ varied from 48.0% in the pooled analysis in China (Zhao et al., 2010) to 82.4% in the meta-analysis in sub-Saharan Africa (Fokom-Domgue et al., 2015), and VIA positivity rates were variable across studies. The specificity of HPV DNA testing for CIN2+ ranged between 84% and 93.6% in all reviews and analyses, except in the meta-analysis in sub-Saharan Africa (73.9%) (Fokom-Domgue et al., 2015); the specificity of VIA for CIN2+ varied from 84% in India (Bobdey et al., 2015) to 90.4% in China (Zhao et al., 2010).

In the pooled analysis of Zhao et al. (2010), a large proportion of participants had received directed biopsies and random biopsies under colposcopy, whereas in the meta-analysis of Fokom-Domgue et al. (2015), colposcopy and directed biopsies performed in all women occurred in only a few of the studies analysed. [Careful consideration is needed when interpreting the accuracy of VIA across different study settings.]

HPV DNA testing has been shown to be superior to VIA as a primary screening technique in detecting cervical neoplasia in postmenopausal women. The study of Holt et al. (2017) found that the sensitivity of HPV DNA testing for both CIN2+ and CIN3+ remained stable near 98%, whereas the corresponding sensitivity of VIA decreased significantly, to 31.0% for CIN2+ and 41.7% for CIN3+.

However, in the study of Arbyn et al. (2008), the pooled sensitivity of HPV DNA testing for CIN2+ was substantially lower than that of VIA (61.9% vs 79.2%), although this difference was not statistically significant (relative sensitivity of HPV vs VIA, 0.883; 95% CI, 0.775–1.007). Several potential explanations for the relatively low sensitivity of HC2 testing have been discussed, including sample contamination or deterioration, limited scope of the hrHPV DNA probe, and misclassification of the outcome, which may result in overestimation of the sensitivity of VIA and underestimation of the sensitivity of HPV DNA testing. Arbyn et al. (2008) reported a relatively high correlation (0.61) between results of VIA and the reference standard (colposcopy), compared with the low correlation (0.13) between results of HC2 testing and colposcopy. [The Working Group noted that VIA and colposcopy were often performed at the same time by health workers who had been trained just before the study began. Potential bias may occur in favour of a test when the test is verified with an imperfect reference standard and results of the two techniques are correlated (e.g. similar inspection after application of acetic acid for both VIA and colposcopy).]

[There is also a potential issue concerning the correlation of reported pooled results, given the overlap between studies being included in different reviews. For example, the study of Sankaranarayanan et al. (2004) has been included in five reviews (Arbyn et al., 2008; Chen et al., 2012; Bobdey et al., 2015, 2016; Fokom-Domgue et al., 2015).] This study was conducted in India and included 18 085 apparently healthy, asymptomatic women aged 25–65 years who were screened with HPV DNA testing, cytology, VIA, and VILI concurrently. The study reported a relatively low sensitivity for both HPV testing

and VIA at some study sites (e.g. in Kolkata, the sensitivity of HPV testing for CIN2/3 was 45.7%, and the sensitivity of VIA was 54.4%). Potential reasons were discussed by the authors, such as the variable expertise of screening providers in specimen collection, unsatisfactory specimens, or DNA losses during HC2 testing (Sankaranarayanan et al., 2004). [The Working Group noted that when studies with such large sample sizes are included, the potential impact on the pooled results in the reviews must be considered.]

(c) Detection rate of cervical neoplasia and cancer with HPV DNA testing versus VIA

Two cluster RCTs in India and South Africa, three cross-sectional studies in China and India, and a pooled analysis of two cohort studies in eastern Europe and Latin America have compared the detection rates of cervical precancer and cancer according to HPV DNA testing and VIA results (Denny et al., 2005, 2010; Sankaranarayanan et al., 2005, 2009; Sarian et al., 2010; Asthana & Labani, 2015; Basu et al., 2015; Zhao et al., 2018). These studies are presented in Table 4.25 and below.

Overall, HPV DNA testing yielded higher detection rates of high-grade cervical lesions compared with VIA.

The RCT conducted in Osmanabad District in India involved 131 746 women aged 30–59 years from October 1999 to November 2003. Clusters, consisting of villages, were randomized into four groups: HPV DNA testing (HC2), VIA, cytology, and a control group that received only health education but no screening at baseline. Immediate colposcopy was offered and directed biopsies were taken from abnormal areas for women in the VIA group. In the other screening groups, colposcopy appointments were made for women who tested positive, and punch biopsy specimens were taken if abnormal findings were present. The HPV testing, VIA, and cytology groups had positivity rates of 10.3%, 13.9%, and 7.0%, respectively, and colposcopy compliance rates of 89.1%, 98.7%, and 87.9%, respectively (Sankaranarayanan et al., 2005, 2009). According to the colposcopy and biopsy findings at baseline, the detection rates were 0.9% for CIN2/3 and 0.3% for cervical cancer in the HPV arm; the detection rates in the VIA arm were similar, at 0.7% for CIN2/3 and 0.3% for cervical cancer.

The other RCT was conducted in South Africa from June 2000 to December 2002. A total of 6555 women aged 35–65 years were recruited, and HPV DNA testing (HC2) was compared with VIA in a screen-and-treat strategy (Denny et al., 2005, 2010). All the participants were screened with HPV DNA testing and VIA at baseline and subsequently randomized to either HPV-and-treat or VIA-and-treat, or to a control group with evaluation delayed for 6 months. Women with a positive test result in both the HPV-and-treat and VIA-and-treat groups underwent cryotherapy. In the HPV DNA testing group, 467 of 2163 women (22%) underwent cryotherapy; in the VIA group, 482 of 2227 women (22%) underwent cryotherapy. At 6 months after randomization, colposcopy was performed by a physician blinded to the group assignment and clinical information for all women. Biopsies were taken for all acetowhite lesions, and appropriate treatment was given for women with CIN2+. At 6 months, the prevalence of CIN2+ was 0.80% (95% CI, 0.40–1.20%) in the HPV-and-treat group, 2.23% (95% CI, 1.57–2.89%) in the VIA-and-treat group, and 3.55% (95% CI, 2.71–4.39%) in the control group. The efficacy of each screen-and-treat approach was presented as the percentage difference in CIN2+ attributable to the approach [(control group – treatment group)/control group]. At the 6-month evaluation, there was a 77% reduction in prevalent CIN2+ in the HPV-and-treat group and a 37% reduction in the VIA-and-treat group compared with the control group. All women with positive HPV DNA or VIA results at enrolment, plus a subset of women who were both HPV DNA-negative and VIA-negative and were

Table 4.25 Detection rates of cervical neoplasia and cancer with HPV DNA testing versus visual inspection with acetic acid (VIA)

Reference Country	Study description	Detection rates for different disease end-points (%) (95% CI), *n*/*N*		Comments
		HPV	VIA	
Denny et al. (2005, 2010) South Africa	RCT design. 6555 unscreened non-pregnant Black women aged 35–65 yr in Khayelitsha, South Africa, were recruited in 2000–2002. All women were screened using HPV DNA test and VIA at baseline, and subsequently randomized to HPV-and-treat (*n* = 2163), VIA-and-treat (*n* = 2227), or control arm (*n* = 2165) with delayed evaluation. All were recalled for colposcopy and biopsy confirmation at 6 mo. In addition, 2708 of them, who were free of CIN2+ at 6 mo, who were HPV DNA-positive or VIA-positive at baseline, plus a subset of women who were both HPV DNA-negative and VIA-negative, were followed up at 12 mo and 36 mo	CIN2+: At 6 mo: 0.80 (0.40–1.20) At 12 mo: 1.42 (0.87–1.97) At 36 mo: 1.50 (NA)	CIN2+: At 6 mo: 2.23 (1.57–2.89) At 12 mo: 2.91 (2.12–3.69) At 36 mo: 3.80 (NA)	Landmark study focusing on HPV DNA testing versus VIA as primary screening methods for screen-and-treat strategy, which fits the situation of low-resource settings. The cumulative detection rates are reported here for each follow-up
Sankaranarayanan et al. (2005, 2009) India	Cluster-RCT design. More than 130 000 healthy women, married but not pregnant, aged 30–59 yr with an intact uterus and no past history of cervical neoplasia, previously unscreened, in rural communities of Osmanabad District, India, were recruited in 1999–2003 and followed up until 2007. Recruited women were randomly assigned to HPV DNA test, VIA, cytology, or control group	CIN2/3: 0.9 (0.6–1.4), 245/27 192 Cervical cancer: 0.2 (0.1–0.4), 73/27 192 CIN2+: 1.2, 318/27 192	CIN2/3: 0.7 (0.3–1.5), 195/26 765 Cervical cancer: 0.3 (0.0–0.7), 82/26 765 CIN2+: 1.0, 277/27 192	Both articles provided the baseline results. Given that Sankaranarayanan et al. (2009) provided more comprehensive information, the main results presented here are based on this article

Table 4.25 (continued)

Reference Country	Study description	Detection rates for different disease end-points (%) (95% CI), *n*/*N*		Comments
		HPV	VIA	
Sarian et al. (2010) Eastern Europe and Latin America	Data were pooled from both the NIS cohort (*n* = 3187) and the LAMS (*n* = 12 114). Women in the NIS cohort attended 6 outpatient clinics in the Russian Federation, Belarus, and Latvia in 1998–2002, and had a mean age of 32.6 yr (range, 15–85 yr). All women underwent Pap testing and HPV DNA testing (HC2). Women in the LAMS cohort had a mean age of 37.9 yr (range, 14–67 yr) and were examined by cytology and VIA, VILI, cervicography, and HPV DNA test (HC2) at 4 clinics in Brazil and Argentina	CIN2+: 2.3, 169/7498	CIN2+: 0.7, 83/12 093	
Asthana & Labani (2015) India	Cross-sectional design. 4658 ever-married women aged 30–59 yr with no history of CIN or cervical cancer, hysterectomy, or the presence of any associated condition were recruited from rural areas in Uttar Pradesh, India, in 2011–2012. All women were screened with HPV DNA test with self-collected sample, HPV DNA test with clinician-collected sample, cytology, and VIA. All screen-positive women were referred for colposcopy and directed biopsy	CIN2+: Self-collected: 2.7 (1.2–4.2) per 1000 women screened Clinician-collected: 3.6 (1.8–5.4) per 1000 women screened CIN3+: Self-collected: 1.5 (0.37–2.6) per 1000 women screened Clinician-collected: 2.4 (0.97–3.8) per 1000 women screened	CIN2+: 1.5 (0.37–2.6) per 1000 women screened CIN3+: 0.21 (−0.21 to 0.63) per 1000 women screened	
Basu et al. (2015) India	Cross-sectional design. 39 740 apparently healthy women aged 30–60 yr from rural districts adjacent to the metropolitan city of Kolkata in eastern India were recruited in 2010–2014. All women were screened with HPV DNA test and VIA	CIN2+: 5.1 per 1000 women screened CIN3+: 3.8 per 1000 women screened	CIN2+: 4.8 per 1000 women screened CIN3+: 2.8 per 1000 women screened	

Table 4.25 (continued)

Reference Country	Study description	Detection rates for different disease end-points (%) (95% CI), *n*/*N*		Comments
		HPV	VIA	
Zhao et al. (2018) China	Cross-sectional study design. 33 823 women aged 35–64 yr, with an intact uterus and with no history of cervical neoplasia or cervical cancer, who were not pregnant and had no suspicious symptoms, and who understood the process and were willing to participate were recruited from rural areas across 7 large geographical regions in China in 2015–2018. In rural areas, women were randomized to initial screening with HPV test (*n* = 15 577), cytology (*n* = 7089), or VIA (*n* = 11 157)	CIN2+: 0.61, 95/15 577	CIN2+: 0.49, 55/11 157	This study is based on real-world data generated from both rural areas (*n* = 33 823) and urban areas (*n* = 30 108) across 7 large geographical regions in China. The results presented here only represent the data from rural areas, because VIA was not applied in urban areas. Women were initially randomized with a 1:1:1 ratio to the 3 arms; however, cytology was not applicable for some rural areas, so VIA was used instead, resulting in more VIA-screened women than HPV-screened and cytology-screened women

CIN, cervical intraepithelial neoplasia; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HC2, Hybrid Capture 2; HPV, human papillomavirus; LAMS, Latin American Screening Study; mo, month or months; NA, not available; NIS, New Independent States; VIA, visual inspection with acetic acid; VILI, visual inspection with Lugol's iodine; yr, year or years.

free of CIN2+ at 6 months were followed up at 12 months and 36 months. At the 12-month follow-up, the cumulative prevalence of CIN2+ was 1.42% (95% CI, 0.87–1.97%) in the HPV-and-treat group, 2.91% (95% CI, 2.12–3.69%) in the VIA-and-treat group, and 5.41% (95% CI, 4.32–6.50%) in the control group in the 2708 women examined. This corresponds to a reduction of 74% in the HPV-and-treat group and of 46% in the VIA-and-treat group compared with the control group (Denny et al., 2005). At the 36-month follow-up, the cumulative detection rate of CIN2+ was lower in the HPV-and-treat group (1.5%) than in the VIA-and-treat group (3.8%), whereas the rate was 3.6% in the control group. This corresponds to a reduction of 72.5% (95% CI, 60.1–85.0%) in CIN2+ in the HPV-and-treat group and a reduction of 32.0% (95% CI, 11.1–52.8%) in CIN2+ in the VIA-and-treat group compared with the control group at 36 months (Denny et al., 2010). In addition, the incidence of CIN2+ detected more than 12 months after enrolment was 0.3% (95% CI, 0.05–1.02%) in the HPV-and-treat group, which was significantly less than in the VIA-and-treat group (1.3%; 95% CI, 0.8–2.1%) and in the control group (1.0%; 95% CI, 0.5–1.7%) ($P = 0.003$) (Denny et al., 2010).

A study involving 33 823 women living in rural areas across seven large geographical regions in China reported detection rates of CIN2+ of 0.61% (95 of 15 577) with HPV DNA testing (*care*HPV, cobas 4800, or Liferiver hrHPV genotyping) and 0.49% (55 of 11 157) with VIA or VILI (Zhao et al., 2018).

In a cross-sectional study in rural India, 4658 eligible women were screened with HPV DNA testing (*care*HPV) with clinician-collected and self-collected samples, VIA, and cytology. For HPV DNA testing with clinician-collected samples, detection rates of CIN2+ were 3.6 (95% CI, 1.8–5.4) per 1000 women screened and detection rates of CIN3+ were 2.4 (95% CI, 0.97–3.8) per 1000 women screened. For HPV DNA testing on self-collected samples, detection rates of CIN2+ were 2.7 (95% CI, 1.2–4.2) per 1000 women screened and detection rates of CIN3+ were 1.5 (95% CI, 0.37–2.6) per 1000 women screened. For VIA, detection rates of CIN2+ were 1.5 (95% CI, 0.37–2.6) per 1000 women screened and detection rates of CIN3+ were 0.21 (95% CI, −0.21 to 0.63) per 1000 women screened (Asthana & Labani, 2015).

A demonstration project in eastern India reported detection rates of CIN2+ of 5.1 per 1000 women screened with HPV DNA testing and 4.8 per 1000 women screened with VIA. For CIN3+, the detection rate with HPV DNA testing (3.8 per 1000 women screened) was significantly higher ($P = 0.016$) than that with VIA (2.8 per 1000 women screened) (Basu et al., 2015).

In a pooled analysis focused on studies in eastern Europe and Latin America, the estimated detection rate of CIN2+ was 2.3% (169 of 7498) in the HPV DNA testing group and 0.7% (83 of 12 093) in the VIA group (Sarian et al., 2010).

(d) Changes in cervical cancer incidence and mortality rates

Only the RCT in Osmanabad District in India has assessed the effect of a single round of HPV DNA testing and VIA as primary screening methods on cervical cancer incidence and mortality rates (Sankaranarayanan et al., 2005, 2009) (Table 4.26). During a follow-up of 8 years, a total of 127 cases of cervical cancer were diagnosed in the HPV DNA testing arm (age-standardized incidence rate [ASIR], 47.4 per 100 000 person-years), compared with 157 cases in the VIA arm (ASIR, 58.7 per 100 000 person-years). A single round of screening with HPV DNA testing also dramatically reduced the incidence of cervical cancer of FIGO stage II or higher compared with VIA screening. The burden of cervical cancer of stage II or higher was reported as 39 cases in the HPV DNA testing arm (ASIR, 14.5 per 100 000 person-years), compared with 86 cases in the VIA arm (ASIR, 32.2 per 100 000 person-years). Fewer cases of cervical cancer

Table 4.26 Age-standardized incidence and mortality rates of cervical cancer with HPV testing versus visual inspection with acetic acid (VIA)

Reference Country	Study description	Age-standardized incidence rate of all cervical cancer (per 100 000 person-years)		No. of cases of cervical cancer of stage II or higher/total no. of cases of cervical cancer (%)		Age-standardized incidence rate of cervical cancer of stage II or higher (per 100 000 person-years)		No. of cases of invasive cervical cancer among screening-negative women/total no. of screening-negative women		Deaths from cervical cancer/ total no. of cases of cervical cancer (%)		Age-standardized mortality rate of cervical cancer (per 100 000 person-years)	
		HPV	VIA	HPV	VIA	HPV	VIA	HPV	VIA	HPV	VIA	HPV	VIA
Sankaranarayanan et al. (2005, 2009) India	See Table 4.25	47.4	58.7	39/127 (30.7%)	86/157 (54.8%)	14.5	32.2	8/24 380 (0.033%)	25/23 032 (0.109%)	34/127 (26.8%)	56/157 (35.7%)	12.7	20.9

HPV, human papillomavirus; VIA, visual inspection with acetic acid.

developed in HPV DNA-negative women (8 cases in 24 380 women; ASIR, 3.7 per 100 000 person-years) than in VIA-negative women (25 cases in 23 032 women; ASIR, 16.0 per 100 000 person-years). Lower cervical cancer-related mortality was also observed in the HPV DNA testing arm. There were 34 deaths in the HPV DNA testing arm (age-standardized mortality rate [ASMR], 12.7 per 100 000 person-years), compared with 56 deaths in the VIA arm (ASMR, 20.9 per 100 000 person-years) (Sankaranarayanan et al., 2009).

(e) Harms

Diagnostic harms can be inferred by the colposcopy referral rates and the PPVs of the screening tests. Details of studies reporting colposcopy referral rates and/or PPVs for HPV DNA testing and VIA are given in Table 4.27. For HPV DNA testing compared with VIA, the different studies did not consistently report a higher or lower proportion of colposcopy referrals or a larger number of colposcopies needed to detect one CIN2+ or CIN3+ case. PPVs were generally higher with HPV DNA testing than with VIA.

4.4.4 Comparison of HPV DNA testing alone versus co-testing

(a) Introduction

Co-testing as a primary screening modality consists of analysing samples for both cytology and HPV at the same time, regardless of the corresponding test result. The analyses can be conducted on the same sample in the case of LBC, where the residual sample can be tested for HPV, or on separate samples taken in sequence at the same visit. The clinical decision about follow-up and/or referral is then made on the basis of the combination of the test results.

The introduction and broader use of LBC since the 2005 *IARC Handbook* has facilitated the use of co-testing in guidelines and routine practice. The technical implementation of co-testing follows the use of cytology and HPV testing as previously described (see Sections 4.3.1 and 4.4.1, respectively). A range of test technologies and analysis platforms exist for both HPV testing and cytology. The interoperability of these sampling methods and platforms enables co-testing but varies across settings and manufacturers.

Studies examining co-testing range from classic RCTs to implementation studies and retrospective analyses of screening test results before precancer and cancer diagnosis. The time perspective for these studies varies: some studies look at the first round of screening results for detection rates and test performance, whereas others present longitudinal evidence for the comparison of cumulative incidence by baseline test results. The early RCTs that compared HPV testing with cytology enabled analyses of co-testing because cytology was done in every participant. In the main results reported by these trials, HPV testing alone was compared with cytology, but the follow-up data provided comparisons between cytology, HPV testing, and co-testing screening strategies (Bulkmans et al., 2004; Naucler et al., 2007; Ronco et al., 2007a; Kitchener et al., 2009a).

In this review, meta-analyses and joint analyses of cohort studies were examined, as well as studies that directly evaluated disease outcomes or test performance of HPV testing alone compared with co-testing as a primary screening modality. Modelling studies, cost–effectiveness analyses, and studies that evaluated co-testing as a follow-up strategy or in conjunction with other biomarkers were excluded. Studies that examined co-testing in specific populations (e.g. non-attenders), as a test of cure, or as a screening programme exit test were also excluded.

Table 4.27 Comparison of potential diagnostic harms of HPV DNA testing versus VIA

Reference	Study description	Colposcopy referrals Referral rate (%) (95% CI), *n*/*N*		PPV for different disease end-points (%) (95% CI), *n*/*N*	
		HPV	VIA	HPV	VIA
Sankaranarayanan et al. (2009)	See Table 4.25	10.3, 2812/27 192	13.9, 3733/26 765	CIN2/3: 11.3, 318/2812 Cancer: 2.6, 73/2812	CIN2/3: 7.4, 277/3733 Cancer: 2.2, 82/3733
Longatto-Filho et al. (2012)	LAMS cohort study. > 12 000 women at 4 clinics in Brazil and Argentina. Large sample size with both cross-sectional and prospective cohorts, which covered regions with different cervical cancer incidence rates. All women were screened with cytology, VIA, VILI, HPV DNA test (HC2) with self-collected sample and clinician-collected sample. Women with a positive screening test result were referred for colposcopy	NA	NA	CIN2+: Self-collected: 9.1 (3.0–22.6) Clinician-collected: 7.9 (6.0–10.1)	CIN2+: 6.1 (4.9–7.6)
Zhao et al. (2013)	START-UP project. 7421 women aged 25–65 yr in 3 counties of China (Yangcheng, Xinmi, and Tonggu) were recruited and tested with *care*HPV, HC2, HPV E6, and VIA using both self-collected and clinician-collected samples. Women with a positive screening test result were referred for colposcopy with directed biopsy. In addition, a randomly selected 10% of women with a negative test result for all the tests also underwent colposcopy	*care*HPV: Self-collected: 14.5 Clinician-collected: 14.4 HC2: Self-collected: 17.9 Clinician-collected: 14.5	7.3	CIN2+: *care*HPV: Self-collected: 11.1 (9.3–13.1) Clinician-collected: 13.0 (11.1–15.2) HC2: Self-collected: 10.0 (8.4–11.7) Clinician-collected: 12.9 (10.9–15.0) CIN3+: *care*HPV: Self-collected: 7.7 (6.2–9.5) Clinician-collected: 9.1 (7.4–10.9) HC2: Self-collected: 6.8 (5.5–8.3) Clinician-collected: 9.0 (7.3–10.8)	CIN2+: 12.7 (10.0–15.9) CIN3+: 9.4 (7.0–12.2)

Table 4.27 (continued)

Reference	Study description	Colposcopy referrals Referral rate (%) (95% CI), *n*/*N*		PPV for different disease end-points (%) (95% CI), *n*/*N*	
		HPV	VIA	HPV	VIA
Asthana & Labani (2015); Labani & Asthana (2016)	See Table 4.25	Self-collected: 2.4 (2.0–2.8), 111/4658 Clinician-collected: 2.9 (2.9–3.4), 136/4658	5.5 (4.9–6.2), 257/4658	CIN2+: Self-collected: 11.7 (6.3–19.1) Clinician-collected: 12.5 (7.4–9.1) CIN3+: Self-collected: 6.3 (2.6–12.6) Clinician-collected: 8.1 (4.1–13.9)	CIN2+: 2.7 (1.1–5.5) CIN3+: 0.4 (0.0–2.2)
Holt et al. (2017)	Postmenopausal women (see Table 4.24 for details)	17.2 (15.9–18.7), 475/2757	6.2 (5.3–7.1), 170/2757	CIN2+: 17.3 (14.1–20.9), 82/475 CIN3+: 9.9 (7.4–12.8), 47/475	CIN2+: 15.3 (10.5–21.3), 26/170 CIN3+: 11.8 (7.5–17.3), 20/170
Wang et al. (2019)	Cross-sectional design. 2668 women aged ≥ 18 yr in Inner Mongolia, China, were screened with HPV DNA test and VIA concurrently. Women with a positive test result were referred for colposcopy	17.5 (16.1–19.0), 467/2668	8.1 (7.1–9.2), 216/2668	CIN2+: 5.6 (3.8–8.0), 26/467	CIN2+: 6.0 (3.6–10.0), 13/216

CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HC2, Hybrid Capture 2; LAMS, Latin American Screening Study; PPV, positive predictive value; START-UP, Screening Technologies to Advance Rapid Testing for Cervical Cancer Prevention–Utility and Program Planning; VIA, visual inspection with acetic acid; VILI, visual inspection with Lugol's iodine; yr, year or years.

(b) Screening performance

A joint database analysis of HPV screening studies included seven studies in six European countries (Dillner et al., 2008) and aimed to estimate the long-term predictive values of HPV-based screening for CIN3+. This analysis included 24 295 women who were screened with HPV testing and cytology at baseline and had at least one additional cervical cytology or histopathology examination during follow-up. The studies differed with respect to the ages of women included, the HPV tests used, and the setting. The cumulative incidence of CIN3+ over 72 months of follow-up was examined by baseline test results, and the test characteristics were reported for cytology, HPV testing, and co-testing with cytology and HPV testing (at least one positive). The cumulative incidence of CIN3+ at 72 months for HPV-negative women was 0.27% (95% CI, 0.12–0.45%), which was similar to that for co-test-negative women at the same time point. At 72 months, the sensitivity of HPV testing for CIN3+ was 90% (95% CI, 80–95%) and the specificity was 88.28% (95% CI, 87.83–88.70%) [recalculated by the Working Group using absolute values without any adjustment; this was erroneously given in the publication]. The corresponding values at 72 months for co-testing with cytology and HPV testing were 92% (95% CI, 84–96%) and 87% (95% CI, 81–93%), respectively.

In a meta-analysis, co-testing with cytology and HC2 testing produced higher detection of CIN2+ (42%; 95% CI, 36–48%) and CIN3+ (33%; 95% CI, 29–37%) compared with cytology alone, and the specificity for the same outcomes was 6% (95% CI, 6–7%) and 8% (95% CI, 7–9%) lower, respectively. When cytology was added to HC2 testing and compared with HPV testing alone, the average sensitivity increased by 5% (95% CI, 4–7%) for CIN2+ and by 2% (95% CI, 1–3%) for CIN3+, and the specificity decreased significantly (ratio for CIN2+, 0.95; 95% CI, 0.94–0.96 and ratio for CIN3+, 0.93; 95% CI, 0.92–0.95). The pooled estimates from the trials showed a non-significant increase in sensitivity for co-testing compared with HPV alone (detection rate ratio for CIN2+, 1.06; 95% CI, 0.97–1.16 and detection rate ratio for CIN3+, 1.04; 95% CI, 0.92–1.17) (Arbyn et al., 2012). [The studies outlined below, which have been conducted since this meta-analysis was completed, used different HPV and cytology platforms but came to broadly the same conclusion.]

(c) Effectiveness

(i) RCTs

RCTs examining the performance of co-testing are outlined in Table 4.28.

Four RCTs in Europe were identified that compared hrHPV co-testing with cytology alone: the NTCC trial in Italy (Ronco et al., 2007a, 2010, 2014), the POBASCAM trial in the Netherlands (Bulkmans et al., 2004; Rijkaart et al., 2012a; Dijkstra et al., 2016), the SwedeScreen trial in Sweden (Naucler et al., 2007; Elfström et al., 2014), and the ARTISTIC trial in the United Kingdom (Kitchener et al., 2009a, b, 2014). The primary results of these trials are reviewed in Section 4.4.2, and long-term follow-up data from these studies have been pooled and provide evidence on the comparison of testing methods and the effectiveness against invasive cervical cancer as an outcome (Arbyn et al., 2012; Ronco et al., 2014).

Both Dijkstra et al. (2016) and Elfström et al. (2014) examined the cumulative incidence of high-grade lesions (CIN2+ or CIN3+). Dijkstra et al. (2016) concluded that the difference between hrHPV testing and hrHPV co-testing with cytology became less pronounced as follow-up time increased, and Elfström et al. (2014) concluded that the difference was minimal over time. Elfström et al. (2014) also calculated the test performance over different follow-up periods (3, 5, 8, and 10 years) and found that although the

Table 4.28 Randomized controlled trials and cohort studies comparing co-testing versus HPV DNA testing

Reference	Study population	Screening exposure Age of included subjects (years)	End-point	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Detection rate	Incidence
Randomized controlled trials							
Mayrand et al. (2007)	10 154 women who sought screening tests for cervical cancer in any of 30 clinics in Montreal and St. John's, Canada	HPV DNA test and cytology 30–69 Pap test result of ASC-US+, or HPV test result of ≥ 1 pg HPV DNA/mL	CIN2+	100.0	92.5	NA	NA
Elfström et al. (2014)	12 527 women who attended the organized cervical screening programme in Sweden. 13-year follow-up of the SwedeScreen RCT of primary HPV screening	HPV DNA test and cytology 32–38	CIN2+	Co-testing: 3-yr: 96.69 (90.25–98.93) 5-yr: 91.22 (84.84–95.07) 8-yr: 82.67 (75.79–87.91) 10-yr: 77.19 (70.16–82.97) HPV testing: 3-yr: 92.23 (84.58–96.25) 5-yr: 86.40 (79.21–91.37) 8-yr: 77.30 (69.95–83.29) 10-yr: 72.45 (65.17–78.71)	Co-testing: 3-yr: 90.32 (89.54–91.05) 5-yr: 90.73 (89.97–91.45) 8-yr: 90.98 (90.22–91.69) 10-yr: 91.10 (90.34–91.81) HPV testing: 3-yr: 94.05 (93.42–94.63) 5-yr: 94.47 (93.85–95.03) 8-yr: 94.69 (94.08–95.24) 10-yr: 94.82 (94.22–95.37)	NA	Cumulative incidence (%) (95% CI) at 13-yr follow-up (no difference between co-testing and HPV testing): CIN2+: 1.63 (1.11–2.32) in the intervention arm CIN3+: 0.84 (0.48–1.47) in the intervention arm

Table 4.28 (continued)

Reference	Study population	Screening exposure Age of included subjects (years)	End-point	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Detection rate	Incidence
Dijkstra et al. (2016)	Of 44 938 women enrolled in the Netherlands, 22 420 were randomized to the intervention group (managed by co-testing results) and 22 518 to the control group (managed only by cytology result)	HPV DNA test and cytology 29–61	CIN3+ and cancer	NA	NA	NA	Incidence ratio (95% CI) (intervention vs control): CIN3+: Cytology-negative and/or HPV-negative: 0.86 (0.63–1.17) Cytology-negative and/or HPV-positive: 0.95 (0.71–1.28) Cytology-positive and/or HPV-negative: 0.62 (0.28–1.37) Cancer: Cytology-negative and/or HPV-negative: 0.58 (0.23–1.48) Cytology-negative and/or HPV-positive: 0.29 (0.10–0.87) Cytology-positive and/or HPV-negative: 5.97 (0.30–119.22)
Han et al. (2020)	182 119 women screened in the primary health-care facilities of 9 districts in Beijing, China, from January 2014 to March 2015	HPV DNA test and cytology 35–64	CIN2+	NA	NA	Co-testing: 5.06 for CIN2+ 1.63 for CIN3+ HPV testing: 3.35 for CIN2+ 2.10 for CIN3+	NA

Table 4.28 (continued)

Reference	Study population	Screening exposure Age of included subjects (years)	End-point	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Detection rate	Incidence
Cohort studies							
Cuzick et al. (2003) Mesher et al. (2010) [6-year follow-up][b]	Multicentre screening study of 11 085 women in the United Kingdom associated with 5 referral centres	HPV test and cytology 30–60	CIN2+	Baseline: Co-testing:[b] 100.0 (96.0–100.0) HPV testing (≥ 2 pg/mL): 96.0 (89.7–98.5)	Baseline: Co-testing:[b] 94.0 (93.4–94.5) HPV testing (≥ 2 pg/mL): 94.4 (93.9–95.0)	NA	6-yr cumulative incidence (%): Co-test-negative: 0.21 HPV-negative: 0.28
Petry et al. (2003)[a]	8466 women attending routine cervical cancer screening in Germany	HPV test and cytology ≥ 29	CIN2+ and CIN3+	CIN2+: Co-testing: 100.0 (93.7–100.0) HPV testing: 97.8 (86.3–99.7) CIN3+: Co-testing: 100.0 (93.7–100.0) HPV testing: 97.3 (83.2–99.6)	CIN2+: Co-testing: 93.8 (91.8–95.3) HPV testing: 95.3 (93.5–96.6) CIN3+: Co-testing: 94.9 (93.1–96.2) HPV testing: 95.2 (93.4–96.5)	NA	NA
Katki et al. (2011)	331 818 women enrolled in co-testing at KPNC starting in 2003–2005 (and with adequate enrolment co-test results) and followed up to 31 December 2009	HPV test and cytology ≥ 30	CIN3+	NA	NA	NA	5-yr cumulative incidence (per 100 000 women per year): Co-test-negative: 3.2 HPV-negative: 3.8

Table 4.28 (continued)

Reference	Study population	Screening exposure Age of included subjects (years)	End-point	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Detection rate	Incidence
Rijkaart et al. (2012b)[c]	VUSA-Screen study. 25 871 women in the Netherlands offered both cytology and hrHPV testing	HPV test and cytology 29–61	CIN2+ and CIN3+	NA[c]	NA[c]	NA	3-yr cumulative risk of CIN2+ (%) (95% CI): Co-test-negative: 0.24 (0.12–0.64) HPV-negative: 0.26 (0.14–0.69) 3-yr cumulative risk of CIN3+ (%) (95% CI): Co-test-negative: 0.05 (0.01–0.42) HPV-negative: 0.06 (0.02–0.46)
Wright et al. (2015)	42 209 women in the USA who underwent cytology and hrHPV testing	HPV test and cytology ≥ 25	CIN3+	NA	NA	NA	3-yr cumulative incidence (%) (95% CI): Co-test-negative: 0.3 (0.1–0.6) HPV-negative: 0.3 (0.1–0.7)
Choi et al. (2016)	922 women who visited the gynaecology clinic at the Korea University Ansan Hospital, Seoul, Republic of Korea, for routine screening or follow-up during an 18-mo period	HPV test and cytology 17–86 (median, 44.7)	CIN2+ and CIN3+	CIN2+: Co-testing: 72.1 HPV testing: 71.3 CIN3+: Co-testing: 59 HPV testing: 61.7	CIN2+: Co-testing: 96.7 HPV testing: 88.1 CIN3+: Co-testing: 100 HPV testing: 98.5	NA	NA

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; hrHPV, high-risk human papillomavirus; KPNC, Kaiser Permanente Northern California; LSIL, low-grade squamous intraepithelial lesion; mo, month or months; NA, not applicable; RCT, randomized controlled trial; yr, year or years.

[a] The follow-up time was not clearly mentioned in the article.

[b] Positive test results defined as cytology ≥ mild (LSIL) or HPV ≥ 2 pg/mL.

[c] Test characteristics for HPV and cytology were reported separately, not as combined test results, and are therefore not noted here.

sensitivity of co-testing was higher than that of HPV testing alone, the specificity was lower for all follow-up periods. In the long-term follow-up of these two trials, the absolute difference in cumulative incidence between co-testing and HPV testing alone remained constant over time and was minimal.

The CCCaST study in Canada randomized 10 154 women aged 30–69 years to either screening with a focus on the HPV testing result or screening with a focus on the cytology result (both tests were performed in both arms). CIN2+ outcomes were reported by screening results (individual and joint HPV and cytology results and HPV genotype-specific results). The test characteristics reported for HPV testing alone and for co-testing with CIN2+ as the outcome were as follows: the sensitivity of HPV testing alone for CIN2+ was 94.6% (95% CI, 84.2–100%) and the specificity was 94.1% (95% CI, 93.4–94.8%) (using a threshold of 1 pg HPV DNA/mL, i.e. 5000 copies of HPV genome per test), and the sensitivity of co-testing for CIN2+ was 100% and the specificity was 92.5%, where the definition of a positive result was ASC-US+ cytology or an HPV test result of 1 pg HPV DNA/mL or above. These estimates were corrected for verification bias and were based on confirmation of the lesion in an excisional specimen (Mayrand et al., 2006, 2007).

In a quasi-RCT implemented in primary health-care facilities, Han et al. (2020) compared cytology with two intervention arms: (i) hrHPV testing alone with cytology triage and (ii) co-testing; the randomization to the intervention arms was done by district. The overall primary outcome was detection rates of CIN2+ by screening strategy; further outcomes included PPV by strategy for CIN2+ and biopsy rates. Detection rates were 5.06‰ for CIN2+ and 1.63‰ for CIN3+ for co-testing, 3.35‰ for CIN2+ and 2.10‰ for CIN3+ for hrHPV testing alone, and 2.47‰ for CIN2+ and 1.24‰ for CIN3+ for cytology. In this study, referral was based on partial genotyping. In the co-testing arm, women who were positive for carcinogenic HPV types other than HPV16 or HPV18 and cytology-negative were referred for repeat testing after 1 year, instead of being deemed negative, as they were in the HPV testing arm.

Taken together, the comparison of co-testing versus HPV DNA testing as examined in these RCTs shows a marginally higher sensitivity for outcomes of CIN2+ and CIN3+ with co-testing than with HPV testing alone. The specificity of co-testing was lower than that of HPV testing alone. The cumulative incidence of high-grade lesions by baseline HPV test-negative women or co-test-negative women showed minor differences over time. Co-test-negative women had a slightly lower cumulative incidence of high-grade lesions, but the difference was not significant (Table 4.28).

(ii) Cohort studies

Cohort studies examining the performance of co-testing are outlined in Table 4.28. They include the Hanover and Tübingen (HAT) study in Germany (Petry et al., 2003), the HART study in the United Kingdom (Cuzick et al., 2003, Mesher et al., 2010), the KPNC cohort in the USA (Katki et al., 2011), and the ATHENA study in the USA (Wright et al., 2015), as well as two studies embedded in routine screening, the VUSA-Screen study in the Netherlands (Rijkaart et al., 2012b) and a study in the Republic of Korea (Choi et al., 2016).

The HAT study included 7908 women aged 30 years and older from routine screening in two cities in Germany in 1998–2000 (Petry et al., 2003). Two samples were taken at baseline; one was analysed with conventional cytology and the other with HPV testing. One round of screening was included, and women were followed up depending on the combination of test results at baseline. Test characteristics were estimated for combinations of baseline test results and the outcomes of CIN2+ and CIN3+. For HPV testing

alone, the sensitivity for CIN2+ was 97.8% (95% CI, 86.3–99.7%) and the specificity was 95.3% (95% CI, 93.5–96.6%). For co-testing (with a cytology threshold of ASC-US+, including unsatisfactory results or any hrHPV positivity), the sensitivity was 100.0% (95% CI, 93.7–100.0%) and the specificity was 93.8% (95% CI, 91.8–95.3%). In the co-testing analysis, positivity in either test resulted in referral. For the outcome of CIN3+, the estimates were similar.

The HART study enrolled 11 085 women aged 30–60 years from routine screening in five cities in the United Kingdom in 1998–2001. As in the HAT study, two samples were taken and analysed with conventional cytology and with HPV testing (Cuzick et al., 2003). Comparisons of the performance of HPV testing alone and co-testing were presented both in the baseline results after one round of screening (Cuzick et al., 2003; test characteristics) and in the long-term follow-up based on an average of 6 years of follow-up (Mesher et al., 2010; cumulative incidence of CIN2+ by baseline test result). At baseline, the sensitivity of HPV testing alone (using a threshold of 2 pg/mL) for CIN2+ was 96.0% (95% CI, 89.7–98.5%) and the specificity was 94.4% (95% CI, 93.9–95.0%), whereas the sensitivity of co-testing, in which the definition of a positive result was mild (similar to LSIL) or worse in cytology or ≥ 2 pg/mL by HPV testing, was 100.0% (95% CI, 96.0–100.0%) and the specificity was 94.0% (95% CI, 93.4–94.5%) (Cuzick et al., 2003). The long-term follow-up of the cohort (Mesher et al., 2010) showed the cumulative incidence of CIN2+ in non-overlapping categories of baseline test results, including HPV-negative women and co-test-negative women; 0.28% of women who were HPV-negative at baseline were diagnosed with CIN2+ during follow-up, and 0.21% of women who were co-test-negative (i.e. HPV-negative and cytology-negative) at baseline developed CIN2+ during follow-up.

KPNC adopted a co-testing strategy in 2003. Data from this large cohort including 331 818 women were reported by Katki et al. (2011) and reflect routine clinical practice. Over 5 years of follow-up, the cumulative incidence of cancer was higher for hrHPV-negative women (3.8 per 100 000 women per year) than for co-test-negative (i.e. hrHPV-negative and cytology-negative) women (3.2 per 100 000 women per year). In a further analysis of the KPNC cohort data (Gage et al., 2014), specific proposed screening strategies in the USA were examined; hrHPV testing alone and co-testing at different intervals were compared with respect to risks of CIN2+, CIN3+, and cancer. The main comparison of interest was the risk of CIN3+ or cancer at 3 years for hrHPV-negative women versus the risk at 5 years for co-test-negative women. The risk of CIN3+ was significantly lower in hrHPV-negative women at 3 years than in co-test-negative women at 5 years (0.069% vs 0.11%; $P < 0.0001$). The risk of cancer was also lower in hrHPV-negative women at 3 years than in co-test-negative women at 5 years (0.011% vs 0.014%), although this difference was not statistically significant. Schiffman et al. (2018) also used the KPNC cohort to examine the relative contribution of the cytology component to co-testing, and concluded that the increased sensitivity of co-testing versus HPV testing alone for detection of treatable precancers and early curable cervical cancers affects very few cases.

In the context of the population-based screening programme in the Netherlands, the VUSA-Screen study (Rijkaart et al., 2012b) examined the effectiveness of co-testing with cervical cytology and hrHPV testing. A total of 25 658 women with adequate baseline samples for cytology and HPV testing were included. Histological results stratified by the baseline screening test result were reported. The 3-year cumulative risk of CIN3+ was 0.06% (95% CI, 0.02–0.46%) for HPV-negative women and 0.05% (95% CI, 0.01–0.42%) for both cytology-negative and hrHPV-negative women. Therefore, adding cytology to hrHPV testing was interpreted to have minimal impact on

evaluating the risk of CIN3+. Test characteristics for hrHPV testing and cytology were reported separately, not as combined test results, and are therefore not given here.

The ATHENA study aimed to evaluate hrHPV testing as a primary screening modality in women aged 25 years or older recruited from routine cervical screening (Wright et al., 2015). The screening strategies examined included hrHPV testing alone (with referral for colposcopy for women who were HPV16- and/or HPV18-positive or ASC-US+ in reflex cytology) and a co-testing strategy that corresponded to United States screening recommendations (cytology alone for women younger than 30 years and co-testing for women aged 30 years or older). The cumulative risks of CIN2+ and CIN3+ were measured over 3 years. The cumulative incidence rate of CIN3+ in HPV-negative women was 0.3% (95% CI, 0.1–0.7%), which was the same as in women who were both HPV-negative and cytology-negative (0.3%; 95% CI, 0.1–0.6%).

In a large cohort trial, the clinical performance of primary HPV screening plus LBC co-testing was compared with that of HPV screening alone and LBC alone at a hospital in Seoul, Republic of Korea, in women aged 17–86 years (Choi et al., 2016). For CIN2+, the sensitivity of primary HPV testing alone was 71.3% and of co-testing was 72.1%; the specificity was 88.1% and 96.7%, respectively. For CIN3+, the sensitivity of HPV testing alone was 61.7% and of co-testing was 59%; the specificity was 98.5% and 100%, respectively.

In recent years, a series of retrospective cohort studies have been conducted that examined the screening history of selected screening cohorts and cohorts of women diagnosed with CIN3+, AIS, or cancer. In a laboratory-based study, Blatt et al. (2015) conducted a retrospective cohort analysis examining the co-test results of 256 648 women aged 30–65 years who had complete results for cytology and HPV testing in 2005–2011 and a follow-up cervical biopsy within 1 year of the index test. Test characteristics for CIN3+ were calculated and reported as follows: the sensitivity of HPV testing alone was 94.0% (95% CI, 93.3–94.7%), and the sensitivity of co-testing was 98.8% (95% CI, 98.6–99.2%). The inclusion criteria required that women had undergone colposcopy and biopsy within 1 year of the index test. By including only women with a follow-up biopsy and limiting the follow-up time to within 1 year, the study excluded a significant percentage of HPV-positive and cytology-negative women who returned for rescreening after more than 1 year; this biased the results in favour of strategies that include cytology at baseline (Castle, 2015; Giorgi Rossi et al., 2016).

Kaufman et al. (2020) took a comparable retrospective approach to analysing co-test results before diagnosis. They examined a total of 13 633 071 co-test results in women aged 30 years or older. Women were included in the analysis if they had at least one LBC and HPV co-test result before a histopathologically confirmed diagnosis of CIN3, AIS, or cancer; 1615 co-tests before 1259 cancer diagnoses and 11 164 co-tests before 8048 CIN3 or AIS diagnoses were included. The results were reported as the proportion of positive results by testing modality before the different diagnoses (cancer was analysed overall and by histopathology), overall and stratified by within 12 months of diagnosis or more than 12 months before diagnosis. In the analysis of test results within 12 months of diagnosis of a cancer, 77.5% of the women were HPV-positive, 85.1% were LBC-positive, and 94.1% were positive on either test. In contrast, the results for more than 12 months before diagnosis show minimal differences between testing modalities. [The focus on test performance within 12 months of a diagnosis presents a significant limitation in the interpretation and application of the results. The authors did not distinguish between screening tests and clinical tests undergone because of symptoms. Tests undergone within a short period of cancer diagnosis often represent tests undergone in the diagnostic workup of a cancer rather than

screening tests; therefore, they are not as indicative of the performance of the testing modality for screening purposes.]

Overall, the performance of HPV testing alone and co-testing in the cohort studies summarized above followed a pattern similar to the results presented in the RCTs: higher sensitivity for co-testing than for HPV testing alone, but lower specificity. The cohort studies presented further data on the risk of high-grade lesions by baseline test result (HPV-negative or co-test-negative). These results confirmed the results of the RCTs and showed little or no difference in cumulative risk between HPV-negative and co-test-negative women over time.

(iii) Harms

In the RCTs reviewed, the PPV for CIN2+ was higher for HPV testing alone than for co-testing. In the long-term follow-up of the SwedeScreen trial, the PPV for CIN2+ was 19.51%, 25.63%, 29.02%, and 31.12% for HPV testing alone at 3, 5, 8, and 10 years, respectively, compared with 13.32%, 17.53%, 20.21%, and 21.56% for co-testing at the same intervals (Elfström et al., 2014). In the CCCaST study, the PPV for CIN2+ was 7.0% for HPV testing alone and 5.5% for co-testing; the colposcopy referral was 6.1% for HPV testing alone and 7.9% for co-testing (Mayrand et al., 2007).

The PPV for HPV testing alone was consistently higher than that for co-testing, although the differences were small. In the joint database analysis of HPV screening studies, the PPV for CIN3+ was 17.1% (95% CI, 12.7–21.4%) for HPV testing alone and 14.7% (95% CI, 9.9–19.0%) for co-testing (Dillner et al., 2008). In the HAT study, the PPV for CIN2+ was 10.9% (95% CI, 8.2–14.2%) for HPV testing alone and 8.6% (95% CI, 6.5–11.3%) for co-testing. The proportion of women referred for colposcopy was 5.2% for HPV testing alone and 6.8% for co-testing (Petry et al., 2003). In the HART study, the PPV for CIN2+ was 15.0% (95% CI, 12.2–18.34%) for HPV testing alone (using a threshold of 2 pg/mL) and 14.4% (95% CI, 11.8–17.5%) for co-testing (using a threshold of mild [similar to LSIL] or worse in cytology or ≥ 2 pg/mL by HPV testing) (Cuzick et al., 2003). In the ATHENA study, there was no significant difference in the PPV for CIN2+ between HPV testing alone (20.2%; 95% CI, 18.3–22.0%) and co-testing (19.5%; 95% CI, 17.6–21.4%) (Wright et al., 2015). The proportion of women referred for colposcopy was higher for co-testing than for HPV testing alone.

4.4.5 HPV testing on self-collected versus clinician-collected samples

(a) Diagnostic accuracy

The diagnostic accuracy of HPV-based testing for detection of CIN2+ and CIN3+ on specimens collected by self-sampling needs to be assessed separately. Clinician-collected cervical specimens have been the reference standard for detection of CIN2+, because exfoliated cells are more likely to be sampled from the target site than with self-sampling, which may include cells from the vagina. Self-sampling is being considered as an alternative to clinician sampling because it is more convenient for women and there are potential cost savings for the health-care system (Campos et al., 2017, 2020). Using a self-sampling device, a woman can collect a sample at home or at a specific collection point; this avoids a speculum examination and leaves the cervix undisturbed, which may improve visual triage of screen-positive women if this is performed on the same day.

Arbyn et al. (2014) evaluated 36 studies, including 154 556 women, on the accuracy of self-collected samples versus clinician-collected samples when used for HPV testing. In the context of screening, HPV testing on self-collected samples detected, on average, 76% (95% CI, 69–82%) of CIN2+ and 84% (95% CI, 72–92%) of CIN3+. The pooled absolute specificity was 86%

Table 4.29 Relative sensitivity and relative specificity of hrHPV assays on self-collected samples versus clinician-collected samples, by sampling device and storage medium[a]

Covariate	Number of studies	Relative sensitivity (95% CI)	Relative specificity (95% CI)
Sampling device			
hrHPV assay based on signal amplification			
Brush	13	0.84 (0.78–0.90)	0.93 (0.91–0.96)
Swab	7	0.85 (0.78–0.91)	0.93 (0.90–0.95)
Lavage	2	0.84 (0.69–1.04)	0.74 (0.55–0.98)
Tampon	1	0.86 (0.78–0.96)	1.02 (1.00–1.03)
hrHPV assay based on polymerase chain reaction			
Brush	12	0.98 (0.95–1.02)	0.95 (0.91–0.99)
Swab	4	0.98 (0.93–1.03)	0.93 (0.89–0.98)
Lavage	4	0.95 (0.87–1.04)	1.09 (0.91–1.30)
Tampon	0	NA	NA
Storage medium			
hrHPV assay based on signal amplification			
Cell-preserving[b]	3	0.84 (0.78–0.90)	0.93 (0.91–0.96)
Virological[b]	15	0.86 (0.81–0.91)	0.95 (0.92–0.98)
Dry samples	0	NA	NA
Other	1	0.90 (0.71–1.13)	0.92 (0.71–1.21)
hrHPV assay based on polymerase chain reaction			
Cell-preserving	6	1.00 (0.96–1.04)	0.92 (0.88–0.97)
Virological	3	0.97 (0.91–1.04)	0.94 (0.89–0.99)
Dry samples	7	0.96 (0.90–1.02)	1.01 (0.94–1.10)
Other	1	0.95 (0.80–1.13)	1.05 (0.69–1.58)

CI, confidence interval; hrHPV, high-risk human papillomavirus; NA, not available.

[a] Relative values were computed using a bivariate normal model, separating studies using an hrHPV assay based on signal amplification or an hrHPV assay based on polymerase chain reaction. Pooling was performed using a bivariate normal model.

[b] When the bivariate model containing covariates did not fit or when the number of studies was < 4, a separate pooling of the relative sensitivity and relative specificity using a model for ratios of proportions was run.

Reproduced with permission from Arbyn et al. (2018).

(95% CI, 83–89%) for CIN2+ and 87% (95% CI, 84–90%) for CIN3+ (Arbyn et al., 2014).

An updated analysis was performed (Arbyn et al., 2018) that included 56 diagnostic accuracy studies up to April 2018 (Table 4.29). Studies were included if the following criteria were met: information was provided on a vaginal sample collected by the woman herself (self-collected sample) followed by a cervical sample collected by a clinician (clinician-collected sample); the same hrHPV assay was performed on both samples; all HPV tests evaluated had been clinically validated according to the Meijer guidelines (Meijer et al., 2009); and the presence or absence of CIN2+ was verified by colposcopy and biopsy in all enrolled women or in women with one or more positive test results. Studies with cytology follow-up for women with negative colposcopy results at baseline assessment were also included but were indexed for sensitivity analyses. Standard methods were used for pooling diagnostic test accuracy (Harbord et al., 2007; Harbord & Whiting, 2009). Indicators included the relative accuracy of tests on self-collected samples versus clinician-collected samples, estimated by incorporating assay category as a covariate in the model. The variation of the accuracy was also evaluated according to the

clinical setting (screening population, high-risk population, follow-up for previous abnormalities, and monitoring after treatment), assay, self-sampling device, and storage medium. [Although the pooled absolute sensitivity and specificity for outcomes CIN2+ and CIN3+ varied by clinical setting, relative values were considered adequate for comparison and were presented first for a screening situation and then for a combination of all clinical settings using only relative indicators.] The relative accuracy of hrHPV assays on self-collected samples versus clinician-collected samples did not vary substantially by clinical setting. The overall relative pooled sensitivity was 0.85 (95% CI, 0.80–0.89) for CIN2+ and 0.86 (95% CI, 0.76–0.98) for CIN3+, and the relative pooled specificity was 0.96 (95% CI, 0.93–0.98) for CIN2+ on self-collected samples versus clinician-collected samples. A higher test positivity and lower PPVs tended to be observed for self-collected samples compared with clinician-collected samples when assays based on signal amplification were used. This was not observed when PCR-based assays were used. PCR-based hrHPV assays were equally sensitive (ratio, 0.99; 95% CI, 0.97–1.02) and slightly less specific (ratio, 0.98; 95% CI, 0.97–0.99) for CIN2+ on self-collected samples versus clinician-collected samples, with similar test positivity and non-significantly lower PPVs.

(b) Additional studies

Since the review by Arbyn et al. (2018), additional studies have been identified that evaluated the accuracy of hrHPV testing for the detection of CIN2+ with vaginal samples and with cervical samples. El-Zein et al. (2018) reported on the Cervical And Self-Sample In Screening (CASSIS) study, which recruited 1217 women aged 21–74 years in Montreal, Canada, attending colposcopy clinics because of an abnormal cytology result. Participants provided three consecutive samples: two different self-collected samples, using the HerSwab device and the cobas 4800 HPV swab, and a clinician-collected sample. The self-collection devices are designed to be anatomically comfortable to enable women to self-collect a sample of exfoliated cervicovaginal cells; the clinician-collected sample was collected with either a swab or a simple brush. [The Working Group did not find the relevant information to confirm whether the clinician collection was performed with a brush or a swab.] The order of the self-sampling devices was assigned randomly. Of 1076 women with complete information (per-protocol population), HPV positivity was high and comparable between the three devices, ranging from 47.4% to 50.5%. Overall, 152 cases of CIN2+ were detected in the per-protocol analysis and 166 in the intention-to-treat analysis.

The relative sensitivity and the relative specificity of self-sampling with the HerSwab device versus clinician sampling for ASC-US+ were 0.94 and 1.07, respectively. The relative sensitivity and the relative specificity of self-sampling with the cobas swab versus clinician sampling for ASC-US+ were 0.94 and 1.02, respectively; the differences were not statistically significant. [The Working Group noted that all women in the study were referred because of an abnormal test result; this may indicate that most women were likely to have a high HPV viral load, and thus the study population may not be suitable for an evaluation of accuracy between tests applied to screening settings.]

In a randomized non-inferiority trial, Polman et al. (2019b) evaluated the diagnostic accuracy of HPV testing on self-collected samples versus clinician-collected samples for the detection of CIN2+ and CIN3+ in a screening population of women aged 29–61 years in the Netherlands. Samples were tested for carcinogenic HPV types using GP5+/6+ PCR EIA. Of the 187 473 women invited to participate, 8212 were randomly allocated to self-sampling first (group A) and 8198 to clinician sampling first (group B) [The response rate was very low,

because self-sampling was an opt-in option of how to be screened.] A total of 7643 women were included in group A and 6282 in group B. A total of 569 (7.4%) self-collected samples and 451 (7.2%) clinician-collected samples tested positive for HPV (RR, 1.04; 95% CI, 0.92–1.17). The sensitivity and specificity of HPV testing for CIN2+ and CIN3+ did not differ between self-collected and clinician-collected samples: for CIN2+, the relative sensitivity was 0.96 (95% CI, 0.90–1.03) and the relative specificity was 1.00 (95% CI, 0.99–1.01), and for CIN3+, the relative sensitivity was 0.99 (95% CI, 0.91–1.08) and the relative specificity was 1.00 (95% CI, 0.99–1.01). [Note that HPV-positive women in both groups were cross-retested with the other collection method, which was done before colposcopy, but the HPV cross-testing results were not disclosed to study participants and were not used for screening management. Although the study had low participation in regular users of screening, the sample size was high in both arms and the study design was powerful.]

In a small cross-sectional study in 104 women aged 25 years or older in Manchester, United Kingdom, attending a colposcopy clinic for management of abnormal cervical screening, Sargent et al. (2019) evaluated the diagnostic accuracy on self-collected vaginal samples and urine and clinician-collected cervical samples for the detection of CIN2+. Vaginal samples and cervical samples were tested using the cobas 4800 and RealTi*me* HPV assays. CIN2+ was detected in 18 women. The sensitivity for detection of CIN2+ was similar for vaginal samples and cervical samples with both HPV assays [relative sensitivity, 1.01] (RealTi*me* assay: 89%, 16 of 18; cobas 4800 assay: 88%, 15 of 17).

(c) Longitudinal evaluation of self-sampling

In the Shanxi Province Cervical Cancer Screening Study I, in China, 1997 non-pregnant women aged 35–45 years with no history of cervical cancer or hysterectomy were enrolled in 1999 via cluster sampling (Zhang et al., 2018). At enrolment, all the women underwent HPV testing on a self-collected sample and a clinician-collected sample. All the women had histologically confirmed results at baseline. HPV testing was done using a signal amplification test (HC2). The relative sensitivities for CIN2+ in clinician-collected samples versus self-collected samples were 1.17 (95% CI, 1.07–1.29) at baseline and 1.15 (95% CI, 1.07–1.25) at 6 years. The values of specificity were identical at baseline and at 6 years (RR, 0.99; 95% CI, 0.97–1.00). Data at 16 years provided similar values.

Issues related to the acceptability of and participation in self-sampling are reviewed in Section 3.3.2.

Aitken et al. (2019) reported on the nationwide implementation of hrHPV-based screening in the Netherlands. In this programme, women receive an invitation to have a cervical sample taken by the provider, but they can also opt for self-sampling at home. Data from the first 18 months of the hrHPV-based screening programme were compared with the previous, cytology-based programme with respect to participation, referral, and detection of CIN. About 8% (36 295 of 454 573) of the women had opted for the use of a self-sampling device. Although no increase in participation could be related to self-sampling, CIN2+ detection was higher in self-collected samples than in clinician-collected samples (1.4% vs 1.1%; $P < 0.001$).

(d) Use of HPV RNA tests on vaginal self-collected samples

The 2018 meta-analysis that assessed the relative accuracy of HPV tests on self-collected versus clinician-collected samples also included three studies in which HPV testing was done with an RNA test (Aptima) (Arbyn et al., 2018). The sensitivity of HPV RNA testing for CIN2+ was significantly lower on self-collected samples than on clinician-collected samples (relative sensitivity, 0.69; 95% CI, 0.52–0.92), whereas the

specificity for CIN2+ was similar in both specimens (relative specificity, 0.97; 95% CI, 0.92–1.02).

Two additional studies evaluated the use of an HPV RNA test (Aptima) on vaginal self-collected samples. Senkomago et al. (2018) studied 350 female sex workers aged 18–50 years in 2009–2013 and compared HPV RNA detection on clinician-collected samples versus self-collected samples. A total of 22 cases with confirmed CIN2+ were detected over a period of 24 months; 18 (82%) were HPV RNA-positive on the clinician-collected samples, and 17 (77%) were HPV RNA-positive on the self-collected samples at baseline [relative sensitivity, 0.85 (95% CI, 0.41–1.76)]. [Note that the referral for biopsy and histological confirmation was done solely on the basis of cytology results, not by HPV test results.] Islam et al. (2020), from the same group, published additional data on HPV RNA testing (Aptima) on dry and wet self-collected samples and found similar performance [the outcome was cytology-confirmed HSIL+].

4.4.6 Comparison of HPV RNA testing versus HPV DNA testing

(a) Use of HPV RNA tests in primary cervical cancer screening

A 2015 review (Arbyn et al., 2015) evaluated the sensitivity and specificity for the detection of CIN2+ and CIN3+ of diverse HPV DNA and RNA assays applied in primary cervical cancer screening and compared them with those of reference HPV DNA tests (HC2 and GP5+/6+ PCR EIA). Six studies that included populations from primary screening were identified that used a 14-HPV type target RNA test (Aptima) and one study that used a 5-HPV type RNA test (PreTect HPV-Proofer). There was no indication that the sensitivity for CIN2+ of the 14-HPV type RNA test was different from that of the comparator HPV DNA test, but it had a higher specificity; the relative sensitivity was 0.98 (95% CI, 0.95–1.01) and the relative specificity was 1.04 (95% CI, 1.02–1.07). The 5-HPV type RNA test was found to be less sensitive but more specific than the comparator HPV DNA test; the relative sensitivity for CIN2+ was 0.74 (95% CI, 0.63–0.88) and the relative specificity was 1.12 (95% CI, 1.10–1.13).

Since that 2015 systematic review, additional studies have been identified that compared the clinical cross-sectional accuracy of an HPV RNA test (Aptima) (Iftner et al., 2015; Maggino et al., 2016; Muangto et al., 2016; Cook et al., 2017) in cervical screening with that of clinically validated hrHPV DNA tests. Other studies aimed to evaluate the longitudinal NPV (Cook et al., 2018; Forslund et al., 2019; Iftner et al., 2019; Zorzi et al., 2020).

In the study of Iftner et al. (2015), 10 040 women aged 30–60 years from the routine cervical cancer screening population of three German centres, in Tübingen, Saarbrücken, and Freiburg, were invited to participate, and 9451 of them were included in the analysis. The study detected 90 cases of CIN2+ and 43 cases of CIN3+. There was no evidence of a difference in the sensitivity for the detection of CIN2+ between the HPV RNA test (Aptima) (87.8%; 95% CI, 80.2–95.5%) and the HPV DNA test (HC2) (93.2%; 95% CI, 87.1–99.2%) [relative sensitivity, 0.94], but the specificity for the detection of CIN2+ of the HPV RNA test was significantly higher than that of the HPV DNA test. For the detection of CIN3+, the sensitivity values were 90.9% for the RNA test and 100.0% for the DNA test [relative sensitivity, 0.90]. For the detection of CIN2+, the specificity values were 96.1% for the RNA test and 94.9% for the DNA test [relative specificity, 1.01]. Women with negative screening test results at baseline were invited to a second round of screening in 2019, and 3295 of them (82.4%) attended follow-up (Iftner et al., 2019). In the second round, 3057 women (92.8%) tested negative by all three screening tests (DNA, RNA, and cytology). A total of 140 women (4.6%) had at least one positive test result at follow-up, and

115 (82%) of those women underwent a colposcopic examination. The 6-year cumulative risks of CIN2+ were 0.62% (95% CI, 0.24–1.59%) for HPV RNA-negative women and 0.47% (95% CI, 0.27–0.81%) for HPV DNA-negative women, and the 6-year cumulative risks of CIN3+ were 0.31% (95% CI, 0.17–0.57%) for HPV RNA-negative women and 0.22% (95% CI, 0.10–0.49%) for HPV DNA-negative women. In women who tested negative by both HPV tests at baseline, the cumulative risk of CIN3+ was 0.17% (95% CI, 0.04–0.75%). The relative sensitivity for the detection of CIN3+ of the HPV RNA test compared with the HPV DNA test was 0.91 [(95% CI, 0.8–1.03)]. [The Working Group noted that the relative risk of CIN3+ between the two cohorts was not provided, and it was estimated to be 1.43, with the 95% confidence interval including unity.]

Cook et al. (2017, 2018) evaluated an HPV RNA test (Aptima) against an HPV DNA test (HC2) within the HPV FOCAL trial. The screening efficacy in women aged 25–65 years of an HPV DNA test (HC2) with LBC triage of all HPV DNA-positive women was compared with LBC screening with HPV DNA triage of women with an ASC-US result. HPV RNA and HPV DNA tests were compared at the baseline screen (3473 women). With HPV DNA as the comparator test, the relative sensitivity of the HPV RNA test for the detection of CIN2+ was 0.96 and for the detection of CIN3+ was 1.00, and the relative specificity was 1.01. In an updated follow-up at 48 months, HPV RNA and HPV DNA tests were compared within the intervention arm (women who tested positive with the HC2 test were triaged with LBC) at baseline and at 48 months for the detection of CIN2+. Women with < CIN2 irrespective of the HPV DNA test result at 48 months were screened with the HPV RNA test, the HPV DNA test, and LBC. At 48 months, 4.8% were HPV RNA-positive and 5.2% were HPV DNA-positive, and the relative sensitivity was close to 1 for both CIN2+ and CIN3+ outcomes. The relative specificity was 1.005. At 48 months, in the 3226 women who were HPV RNA-negative at baseline, 12 of 2858 (0.4%) had CIN2+; in the 3184 women who were HPV DNA-negative at baseline, 13 of 2821 (0.5%) had CIN2+. There was no difference in the detection of CIN2+ at 48 months between the HPV RNA-negative and HPV DNA-negative women at baseline, and accuracy estimates at 48 months were similar.

Forslund et al. (2019) studied a population-based cohort of 95 023 women in Sweden with available cervical samples collected between May 2007 and January 2012 and frozen at −80 °C. Registry linkages identified that 1204 of these women had CIN3+ after 4 months to 7 years since enrolment. Baseline samples were analysed with an HPV RNA test (Aptima) and an HPV DNA test (cobas 4800), and results from both tests were obtained for 1172 women. Both for women younger than 30 years and for women aged 30 years or older, the HPV RNA and HPV DNA tests had similar sensitivities for the detection of CIN3+. In women aged 30 years or older, the longitudinal sensitivities for CIN3+ occurring during the 2-year period 5–7 years after enrolment were lower for the HPV RNA test, with a relative sensitivity of 0.92 and a relative longitudinal NPV of 1.

Maggino et al. (2016) and Zorzi et al. (2020) published the baseline data and the 5-year follow-up data for two cohorts in two neighbouring areas in Italy, one tested with an HPV RNA test (Aptima) and the other with an HPV DNA test (HC2). Women in both cohorts who tested negative at baseline (22 338 women in the RNA cohort and 68 695 women in the DNA cohort) were followed up. The study reports on the 5-year risk of CIN2+ and CIN3+ and the performance parameters at the 3-year rescreening of a negative HPV RNA test compared with those of a negative HPV DNA test in the two cohorts. The Veneto Cancer Registry was checked to search for invasive cancers and CIN3 diagnosed

up to 5 years after the negative baseline test. The baseline data showed that the proportion of positive Pap tests in HPV-positive women and the cumulative referral rate for colposcopy were both higher (52.8% vs 38.2%, $P < 0.0001$; 4.8% vs 4.5%, $P = 0.04$) in the HPV RNA cohort than in the HPV DNA cohort. The ratio of positive HPV tests, of referral for colposcopy, and of detection of CIN2+ in the RNA cohort compared with the DNA cohort were as follows: HPV prevalence ratio, 1.08 (95% CI, 0.99–1.17); referral ratio, 1.06 (95% CI, 0.95–1.18); and CIN2+ detection ratio, 0.85 (95% CI, 0.54–1.33). The relative 5-year cumulative risks of CIN2+ in the RNA cohort and the DNA cohort were 1.1 and 1.5 per 1000 women, respectively (ratio, 0.74; 95% CI, 0.45–1.16), and the risks of cancer were 4.5 and 8.7 per 100 000 women, respectively (ratio, 0.51; 95% CI, 0.01–4.22). [The study has a major caveat, because the comparison was not performed within the same study population but compared two cohorts in parallel.]

[An important issue relating to HPV RNA tests has been the difficulty of estimating the length of time for which a baseline test has negative predictive value. Given the overall slightly lower sensitivity of the HPV RNA tests, the safety of intervals between screening rounds of longer than 5 years remains uncertain. The studies reporting on longer than 5 years are those of Iftner et al. (2019) and Forslund et al. (2019), who reported on women with negative results at baseline. Although Iftner et al. (2019) did not detect a statistically significant difference between HPV RNA tests and HPV DNA tests, Forslund et al. (2019) found a higher longitudinal sensitivity for the HPV DNA test that was evaluated. The lower sensitivity of HPV RNA tests applied in screening settings may affect the longitudinal NPV at 5 years.]

(b) Use of HPV RNA tests in triage of women with minor abnormal cervical cytology

Ovestad et al. (2011) evaluated two HPV RNA tests – a 5-HPV type RNA test (PreTect HPV-Proofer) and a 14-HPV type RNA test (Aptima) – and two HPV DNA tests – Amplicor and cobas 4800 – for the triage of women with ASC-US or LSIL cytology results. The study included 528 women in Norway selected from a consecutive population-based follow-up of LBC samples for the diagnosis of CIN2/3. [The study has several limitations. One is that the population is a referral population for abnormal results and may not be the most suitable to compare screening tests with a lower HPV viral load. Furthermore, the two RNA tests that were evaluated targeted different sets of HPV types. The 14-HPV type RNA test was significantly more specific than the Amplicor DNA test (ratio, 2.14; 95% CI, 1.23–2.73) and was more sensitive than the 5-HPV type RNA test (ratio, 1.91; 95% CI, 1.43–2.56) but less specific (ratio, 0.47; 95% CI, 0.34–0.63).]

Arbyn et al. (2013b) performed a meta-analysis of studies reporting on an HPV RNA test (Aptima) compared with an HPV DNA test (HC2) for the triage of women with ASC-US or LSIL cytology results. Eight studies were retrieved, which included 1839 ASC-US cases and 1887 LSIL cases. The outcome was histological detection of CIN2+ or CIN3+. All of the women included had undergone a colposcopic evaluation (this may not imply that all of the women had had a biopsy); a negative colposcopy was considered as ascertainment for the absence of disease when no biopsies were taken. Table 4.30 summarizes the relative accuracy of the HPV RNA test compared with the HPV DNA test for CIN2+ or CIN3+ at a threshold of abnormal cytology of ASC-US or LSIL. The sensitivity of the HPV RNA test was not significantly different from that of the HPV DNA test for either of the outcomes measured, but the specificity of the

Table 4.30 Pooled relative sensitivity and specificity of HPV RNA testing compared with HPV DNA testing

Baseline outcome	Outcome after triage	Parameter	Ratio (HPV RNA/HPV DNA) (95% CI)
ASC-US	CIN2+	Sensitivity	1.01 (0.97–1.06)
ASC-US	CIN2+	Specificity	1.19 (1.08–1.31)
ASC-US	CIN3+	Sensitivity	1.01 (0.96–1.06)
ASC-US	CIN3+	Specificity	1.18 (1.08–1.29)
LSIL	CIN2+	Sensitivity	0.96 (0.92–1.03)
LSIL	CIN2+	Specificity	1.37 (1.22–1.54)
LSIL	CIN3+	Sensitivity[a]	0.98 (0.91–1.06)
LSIL	CIN3+	Specificity[a]	1.35 (1.11–1.66)

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion.

[a] The SAS macro MetaDAS failed to converge. Therefore, the pooled relative sensitivity and specificity were computed separately as ratios.

HPV RNA test was significantly higher both for CIN2+ and for CIN3+. [The study is robust, because the overall analysis was not heterogeneous and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) evaluation did not identify major issues.]

The meta-analysis of Verdoodt et al. (2013) compared the diagnostic accuracy of two HPV RNA tests (PreTect HPV-Proofer and NucliSENS EasyQ), both of which target five HPV types, with that of an HPV DNA test (HC2) for the detection of CIN2+ and CIN3+ in women with ASC-US or LSIL. In women with ASC-US or LSIL, HPV RNA testing was significantly more specific than HPV DNA testing for the detection of CIN2+ (ratio 1.98; 95% CI, 1.7–2.3) or CIN3+ (ratio, 3.36; 95% CI, 2.82–4.0), but was significantly less sensitive for the detection of CIN2+ (ratio, 0.80; 95% CI, 0.73–0.87) and CIN3+ (ratio, 0.74; 95% CI, 0.69–0.80). [The comparison between the HPV RNA tests and the HPV DNA test is expected to be limited because of the difference in the HPV types targeted; the HC2 test targets 13 hrHPV types, whereas both RNA tests that were evaluated target five hrHPV types.]

As a part of the Clinical Evaluation of Aptima mRNA (CLEAR) study, Stoler et al. (2013) evaluated HPV RNA testing for the triage of 939 women with ASC-US cytology for colposcopy referral. A cervical specimen in liquid cytology medium was used to test in a blinded fashion for HPV DNA (cobas 4800), for HPV RNA (Aptima), and for RNA type-specific HPV16, HPV18, and HPV45 for those samples that were HPV RNA-positive. The final diagnoses were based on a consensus panel review of the histology of the biopsy specimen. For detection of CIN2+, the HPV RNA test and the HPV DNA test were equally sensitive (ratio, 1.0; 95% CI, 0.91–1.10), and the HPV RNA test was more specific than the HPV DNA test (ratio, 1.13; 95% CI, 1.04–1.21). Risk stratification using partial HPV genotyping was similar for the two assays. [The CLEAR study had been included in the previous meta-analysis by Gen-Probe (2011), in which data were extracted from a report published by the United States FDA.]

Cook et al. (2017) evaluated an HPV RNA test (Aptima) against an HPV DNA test (HC2) within the HPV FOCAL trial (described above). In addition to the main strategy, further triage strategies to refer women for colposcopy were compared in HPV DNA-positive or HPV RNA-positive women as follows:

Table 4.31 Colposcopy referral rates and CIN2+ and CIN3+ detection rates by baseline and triage strategies

Primary test result	Triage strategy result	Number of women screened	Colposcopy referral rate (%) (95% CI)	Detection rate (per 1000 women screened) (95% CI)	
				CIN2+	CIN3+
Baseline HPV DNA+	ASC-US+	125	36.0 (30.3–42.7)	11.2 (8.2–15.3)	4.0 (2.4–6.8)
Baseline HPV DNA+	Persistent HPV DNA+ and/or ASC-US+	86	24.8 (20.1–30.5)	3.2 (1.8–5.7)	1.2 (0.4–3.0)
Baseline HPV RNA+	ASC-US+	107	30.8 (25.6–37.1)	10.9 (8.0–15.0)	4.0 (2.4–6.8)
Baseline HPV RNA+	HPV16/18/45+	67	19.3 (15.2–24.4)	7.8 (5.4–11.3)	3.5 (2.0–6.0)
Baseline HPV RNA+	ASC-US+, or NILM and HPV16/18/45+	133	38.3 (32.4–45.2)	12.4 (9.2–16.6)	4.6 (2.8–7.5)

ASC-US, atypical squamous cells of undetermined significance; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; NILM, negative for intraepithelial lesion or malignancy.

(i) HPV DNA-positive and ASC-US+, (ii) HPV DNA-positive with 12-month HPV persistence and/or ASC-US+, (iii) HPV RNA-positive and ASC-US+, (iv) HPV RNA-positive and HPV16/18/45-positive, and (v) HPV RNA-positive and ASC-US+, or HPV RNA-positive and NILM and HPV16/18/45-positive. [Genotyping was performed with an HPV RNA (Aptima) HPV16/18/45 genotyping assay.] Table 4.31 shows the accuracy results of the different triage strategies. [The Working Group noted that women who were HPV DNA-negative but HPV RNA-positive were not referred for colposcopy; this could lead to an underestimate of an added value of the HPV RNA test, although this should be minimal, given the slightly lower sensitivity of HPV RNA tests compared with HPV DNA tests.] Compared with the triage strategy of immediate referral for colposcopy of women who were HPV DNA-positive with abnormal cytology at baseline and those with 12-month HPV persistence (60.8 per 1000 women screened), the colposcopy referral rate was significantly lower (38.3 per 1000 women screened; $P < 0.001$) in the strategy in which HPV RNA-positive women with abnormal LBC or HPV16/18/45 positivity were referred at baseline.

4.4.7 Triage of women with a positive primary HPV screening test result

Testing for the presence of HPV (in the absence of triage) is inherently limited in terms of its specificity for the presence of histologically confirmed CIN2+ and CIN3+ (Arbyn et al., 2012). Although hrHPV positivity predicts an increased risk of the future development of CIN2+ and CIN3+ (even if disease is not present at the time of the index screening test) (Katki et al., 2011), the lower cross-sectional specificity nevertheless implies that some screen-positive women might be followed up unnecessarily. Therefore, appropriate triage testing, management, and follow-up of HPV-positive women is of critical importance to optimize the balance of benefits and harms of primary HPV screening. The general principle is to refer for diagnostic workup women who are at a higher risk of having a current or incipient precancer, to return to routine screening women who are at low risk, and to keep under surveillance women who are at intermediate risk (Arbyn et al., 2017).

(a) Methods

For this *Handbook*, the Working Group updated a previous meta-analysis on the accuracy of six tests or combinations of tests used to triage hrHPV-positive women identified at screening for the detection of underlying cervical precancer (HAS, 2019). Literature retrieval was extended up to 31 January 2020. The Working Group drafted the review question in PICOS form (population, intervention, comparator, outcome, and studies) to determine the inclusion and exclusion criteria for the studies. PICOS components of the research question are summarized in Box S1 (Annex 1; web only; available from https://publications.iarc.fr/604). Studies were eligible if (i) cross-sectional and/or longitudinal outcome data were available for women with a positive hrHPV screening test result triaged with an index test, and (ii) verification with the reference standard (colposcopy and targeted biopsy, possibly complemented with random biopsies and/or endocervical curettage) was performed on all women or on women with at least one positive triage test result. Normal satisfactory colposcopy without biopsy was accepted as ascertainment of the absence of CIN2+. The methodological quality of the selected studies was assessed using the QUADAS-2 checklist (Whiting et al., 2011).

The current review was limited to one-time (reflex) triage strategies for women with a positive hrHPV test result on a clinician-collected cervical specimen using the following tests: (i) cytology at a threshold of ASC-US+, (ii) genotyping for HPV16/18, (iii) p16/Ki-67 immunocytochemistry (dual staining), (iv) VIA, (v) the combination of HPV16/18 genotyping and cytology, and (vi) the combination of HPV16/18 genotyping and VIA. Strategies involving other triage tests or combinations and two-time triage strategies (including surveillance of women who were reflex triage-negative) and triage of women with an HPV-positive self-collected sample are not included here.

The numbers of true positives and false positives and true negatives and false negatives were extracted from each primary study to compute the sensitivity, specificity, PPV, NPV, the complement of NPV (i.e. 1 – NPV [cNPV]), the test positivity rate, and the underlying prevalence of CIN2+ and CIN3+. Standard statistical procedures for pooling diagnostic accuracy data were used (Leeflang et al., 2008). The results were displayed graphically in forest plots and summary ROC (sROC) curves. For each triage approach, the relative sensitivity and specificity compared with reflex cytology at a threshold of ASC-US+ was also assessed. Finally, to illustrate the principle of triage as it applies in a specific local setting, the implied performance of CIN3+ risk-based stratification was considered for each triage approach, given examples of potentially acceptable local risk thresholds for either return to routine screening or referral for colposcopy. The numbers of false-positive and true-positive and false-negative and true-negative results were calculated for a population of 1000 triaged hrHPV-positive women, as were the PPV and cNPV for CIN3+. In addition, the proportion of triage-positive women who would be referred for colposcopy was calculated, together with the number of women who must be referred for colposcopy to detect one case of CIN3+ (= 1/PPV). For this exercise, three background situations were simulated in terms of the underlying risk of CIN3+: (i) a low-risk situation, with a prevalence of CIN3+ of 5% (corresponding to the 10th percentile of the distribution of observed prevalence throughout the meta-analysis); (ii) an intermediate-risk situation, with a prevalence of CIN3+ of 8% (corresponding to the median prevalence); and (iii) a high-risk situation, with a prevalence of CIN3+ of 17% (corresponding to the 90th percentile of the distribution of observed prevalence throughout the meta-analysis).

(b) Results

Overall, 93 studies were included in the meta-analysis; the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram is shown in Fig. S1 (Annex 1; web only; available from https://publications.iarc.fr/604). Most QUADAS-2 items for the included studies were assessed as satisfactory or borderline; see Fig. S2 (Annex 1; web only; available from https://publications.iarc.fr/604). The summary results of all the meta-analyses are presented in Table 4.31. The detailed results are presented in Figs. S3–S5, and Table S1 (Annex 1; web only; available from https://publications.iarc.fr/604).

(i) Triage with cytology at a threshold of ASC-US+

The pooled sensitivity for CIN2+ in 39 studies was 72% (95% CI, 65–77%) and for CIN3+ in 28 studies was 78% (95% CI, 69–84%), and the pooled specificity for < CIN2 was 75% (95% CI, 69–80%) (see Fig. 4.5, Fig. S3 [Annex 1; web only; available from https://publications.iarc.fr/604], Table 4.32). The pooled relative sensitivity for the detection of CIN2+ was higher (ratio, 1.22; 95% CI, 1.04–1.44) and the specificity was lower (ratio, 0.75; 95% CI, 0.64–0.88) in the group of studies in which the cytologists were aware of the HPV status compared with the group of studies in which the cytologists were blinded to the HPV status; for the sROC curves stratified by the cytologists' knowledge of the HPV status, see Fig. S4B (Annex 1; web only; available from https://publications.iarc.fr/604). For the detection of CIN3+, the impact of the cytologists' knowledge of the HPV status was smaller (detailed results not shown). There were no significant differences in accuracy for detection of CIN2+ or CIN3+ between conventional cytology and LBC methods used in triage of HPV-positive women when ASC-US+ was used as the threshold (detailed results not shown). However, the accuracy of cytology at a threshold of ASC-US+ for CIN2+ was higher in HPV16/18-positive women than in HPV16/18-negative women (detailed results not shown).

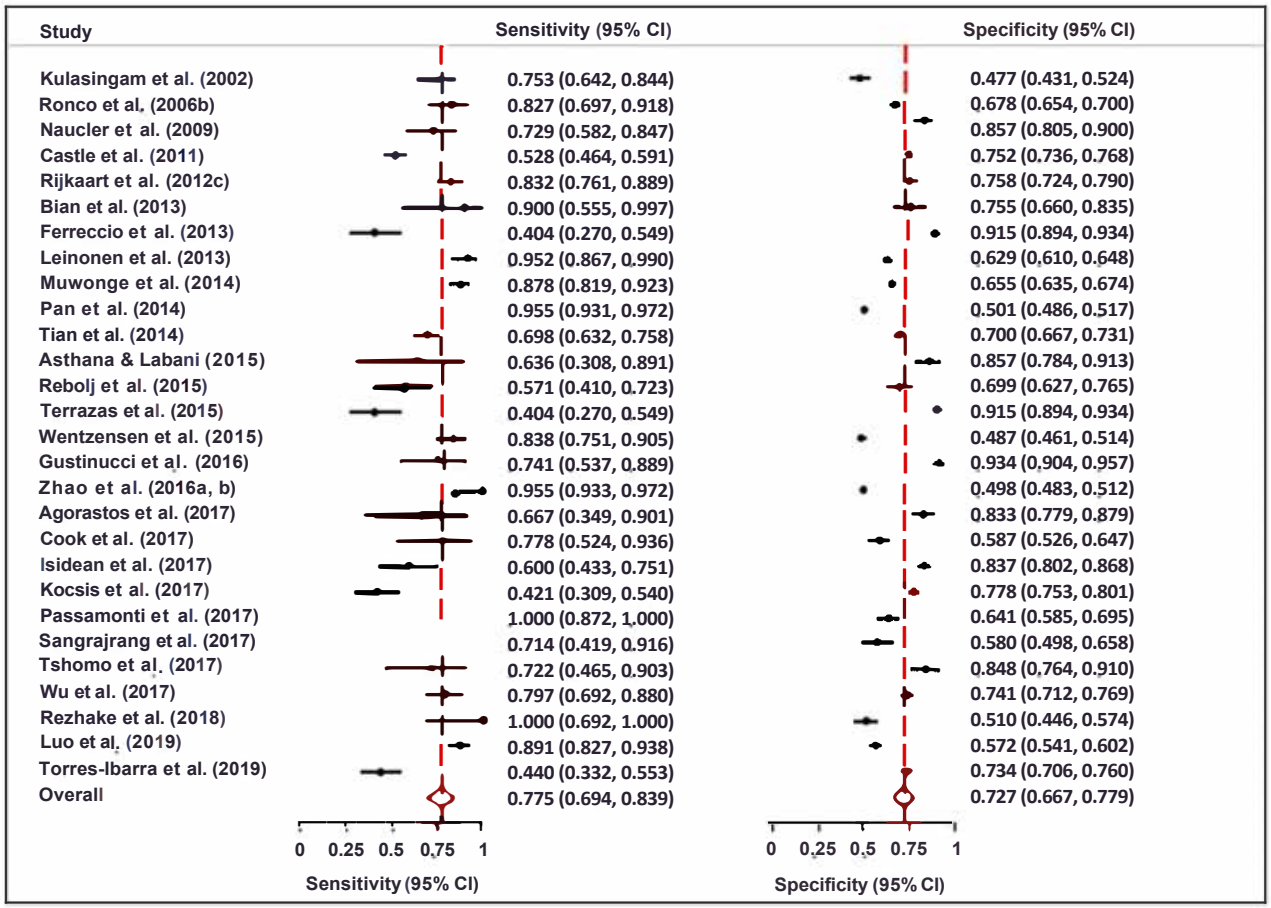
(ii) Triage with VIA

Fig. 4.6 shows a forest plot for the meta-analysis of the absolute sensitivity and specificity of triage of hrHPV-positive women with VIA for the detection of CIN3+. The sensitivity was extremely heterogeneous between studies, varying from 6% (Asthana & Labani, 2015) to 100% (Almonte et al., 2020) for CIN2+ and from 7% to 100% for CIN3+ (Fig. 4.6). Exclusion of these two extreme observations yielded a pooled sensitivity of 64% (95% CI, 56–72%) for CIN2+ and of 69% (95% CI, 61–75%) for CIN3+, and a pooled specificity for < CIN2 of 79% (95% CI, 73–84%) (Table 4.32). The relative accuracy estimates (VIA compared with cytology) did not differ from unity; the sensitivity ratio was 1.15 (95% CI, 0.76–1.83) for CIN2+ and 1.01 (95% CI, 0.70–1.45) for CIN3+, and the specificity ratio for < CIN2 was 0.82 (95% CI, 0.58–1.16) (detailed results not shown). Very wide interstudy variation in the relative sensitivity and specificity was observed (I^2 > 97%; data not shown).

(iii) Triage with HPV16/18 genotyping

The pooled sensitivity of HPV16/18 genotyping to triage hrHPV-positive women was 53% (95% CI, 50–56%) for CIN2+ and 61% (95% CI, 57–65%) for CIN3+, and the pooled specificity for < CIN2 was 75% (95% CI, 70–79%) (Table 4.32, Fig. S4 and Fig. S5 [Annex 1; web only; available from https://publications.iarc.fr/604]). For the detection of CIN2+, HPV16/18 genotyping was less sensitive (ratio, 0.85; 95% CI, 0.75–0.96) but similarly specific (ratio, 1.03; 95% CI, 0.95–1.12) compared with cytology at a threshold of ASC-US+. For the detection of CIN3+, there was no significant difference in accuracy between triage with HPV16/18 genotyping and reflex cytology at a threshold of ASC-US+.

Fig. 4.5 Meta-analysis of the absolute sensitivity and specificity of triage of hrHPV-positive women with reflex cytology at a threshold of ASC-US+ for the detection of CIN3+



ASC-US+, atypical squamous cells of undetermined significance or worse; CI, confidence interval; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; hrHPV, high-risk human papillomavirus.
Created by the Working Group.

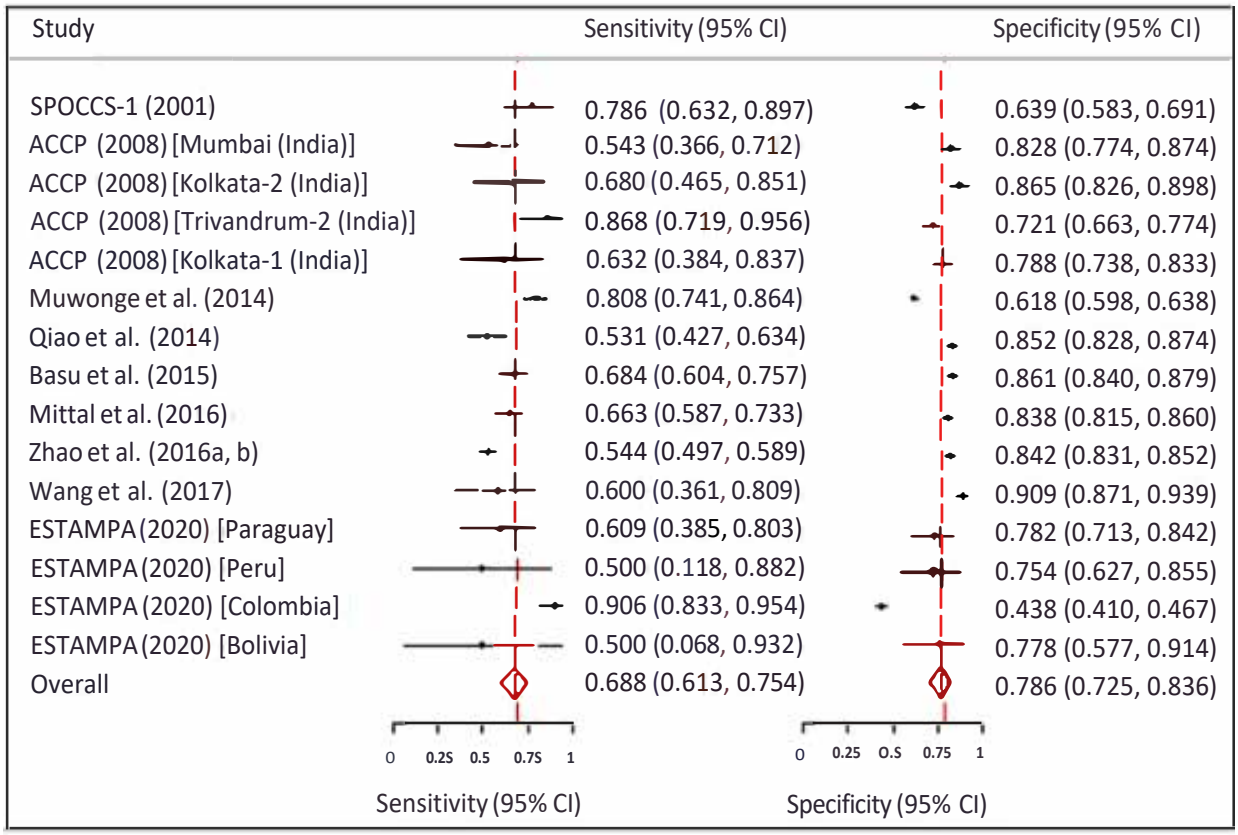
(iv) Triage with immunocytochemistry (dual staining) for p16/Ki-67

Dual staining for p16/Ki-67 was more sensitive than reflex cytology at a threshold of ASC-US+, but the difference was significant only for CIN2+ (81% vs 72%; ratio, 1.12; 95% CI, 1.01–1.25) and not for CIN3+ (Table 4.32, Fig. S4 and Fig. S5 [Annex 1; web only; available from https://publications.iarc.fr/604]). The specificity of dual staining for < CIN2 was similar to that of cytology at a threshold of ASC-US+ (69% vs 75%).

(v) Triage with HPV16/18 genotyping combined with cytology or VIA

HPV16/18 genotyping is usually not used as a stand-alone method to triage hrHPV-positive women. A combined strategy in which HPV16/18-positive women are directly referred for colposcopy and women who are positive only for other carcinogenic HPV types are further triaged with cytology, with referral for colposcopy when cytology shows ASC-US+, had a sensitivity of 83% (95% CI, 79–86%) for CIN2+ and 86% (95% CI, 72–84%) for CIN3+, and the specificity

Fig. 4.6 Meta-analysis of the absolute sensitivity and specificity of triage of hrHPV-positive women with VIA for the detection of CIN3+



ACCP, Alliance for Cervical Cancer Prevention; CI, confidence interval; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; hrHPV, high-risk human papillomavirus; SPOCCS, Shanxi Province Cervical Cancer Screening Study; VIA, visual inspection with acetic acid.
ACCP (2008): main reference is Arbyn et al. (2008); ESTAMPA (2020): main reference is Almonte et al. (2020); SPOCCS-1 (2001): main reference is Belinson et al. (2001).
Note: Unpublished data were provided by IARC from the ESTAMPA study (Almonte et al., 2020).
Created by the Working Group.

for < CIN2 was 55% (95% CI, 48–62%). Only two studies provided data for the combination of HPV16/18 genotyping and VIA (Table 4.32, Fig. S4 and Fig. S5; Annex 1; web only; available from https://publications.iarc.fr/604).

(vi) Utility of triage based on the post-test risk of CIN3+

Fig. 4.7 is an example pre-test–post-test probability plot showing the risk of CIN3+ through the triage pathway applied to hrHPV-positive women starting with partial genotyping (i.e. HPV16/18-positive). Women who are positive only for other hrHPV types receive a secondary triage with cytology at a threshold of ASC-US+. In Fig. 4.7, a median underlying risk (8%) of CIN3+ in hrHPV-positive women (notionally representing, for example, a population in a middle-income or high-income country) is assumed. Triage with HPV16/18 genotyping enables post-test separation of the population of women into those who are positive for HPV16/18, with a higher risk (almost 20%) of CIN3+, and those who are negative for HPV16/18, with a

Table 4.32 Pooled cross-sectional sensitivity and specificity of selected tests used to triage hrHPV-positive women to detect CIN2+ or CIN3+

Triage test	Outcome	Number of studies	Referral rate (%) (IQR or range)[a]	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
ASC-US cytology (all)	CIN2+	39	33.8 (28.9–43.8)[a]	71.5 (65.2–77.1)	74.7 (69.2–79.5)
VIA	CIN2+	17	22.4 (19.3–35.3)[a]	64.2 (56.1–71.5)	79.2 (73.0–84.2)
HPV16/18 genotyping	CIN2+	16	30.7 (20.2–34.3)[a]	52.9 (50.2–55.7)	74.9 (70.3–79.0)
p16/Ki-67 dual staining	CIN2+	5	36.5 (29.4–46.0)	80.8 (74.5–85.8)	69.0 (61.1–75.9)
HPV16/18 genotyping, ASC-US+ cytology if positive for other hrHPV types	CIN2+	12	53.5 (44.6–68.8)[a]	82.6 (79.2–85.5)	55.4 (48.2–62.4)
HPV16/18 genotyping, VIA if positive for other hrHPV types	CIN2+	2	45.3 (43.3–49.4)	87.2 (78.4–92.8)	59.9 (56.2–63.4)
ASC-US cytology (all)	CIN3+	28	[b]	77.5 (69.4–83.9)	72.7 (66.7–77.9)
VIA	CIN3+	15	[b]	68.8 (61.3–75.4)	78.6 (72.5–83.6)
HPV16/18 genotyping	CIN3+	10	[b]	61.2 (57.2–65.2)	74.9 (68.7–80.2)
p16/Ki-67 dual staining	CIN3+	4	[b]	85.1 (77.4–90.5)	63.8 (55.6–71.2)
HPV16/18 genotyping, ASC-US+ cytology if positive for other hrHPV types	CIN3+	9	[b]	85.8 (72.1–84.2)	67.5 (60.1–72.4)
HPV16/18 genotyping, VIA if positive for other hrHPV types	CIN3+	2	[b]	91.5 (79.4–96.8)	57.6 (54.0–61.0)

ASC-US+, atypical squamous cells of undetermined significance or worse; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; HPV, human papillomavirus; hrHPV, high-risk human papillomavirus; IQR, interquartile range; VIA, visual inspection with acetic acid.

[a] Referral rate is the percentage of hrHPV-positive women with a positive triage test result. IQR if ≥ 8 studies; range if < 8 studies.

[b] Referral rate is not given for the CIN3+ outcome, because it should be the same as for the CIN2+ outcome.

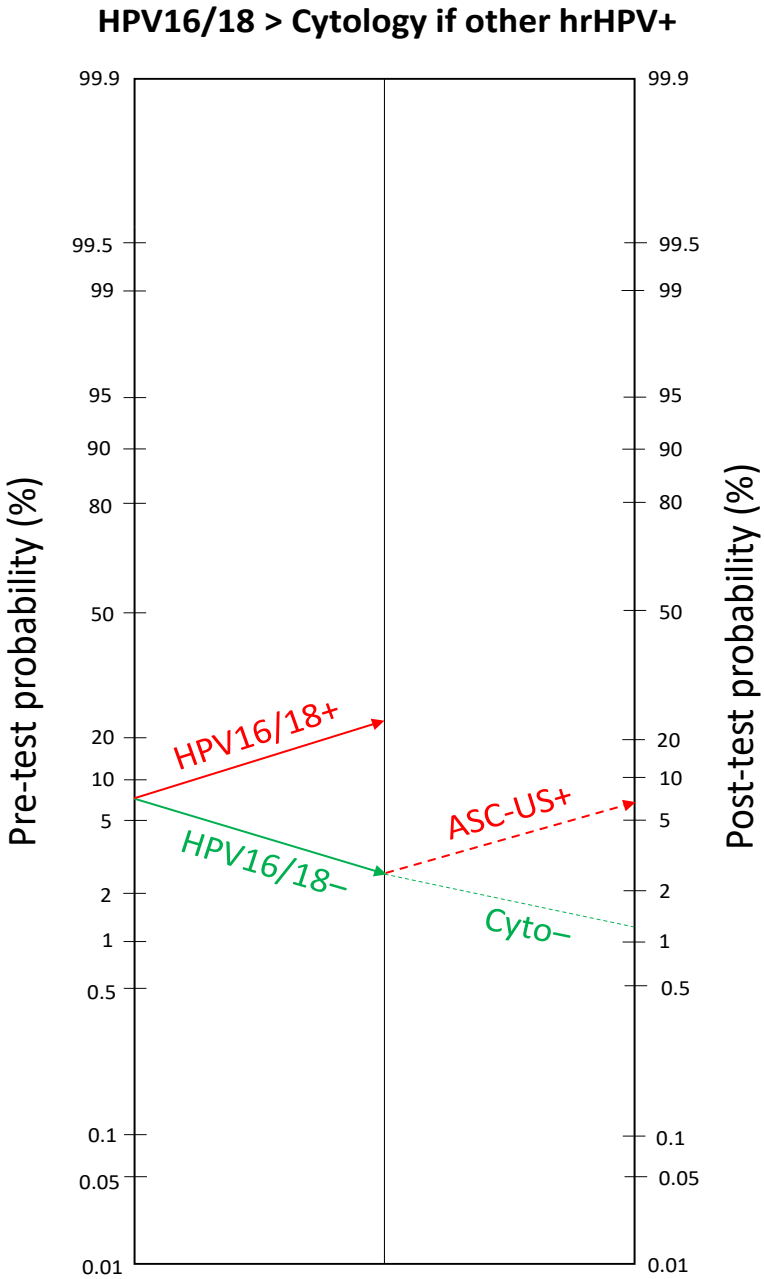
lower risk (about 3%). This latter group can be further triaged with cytology to resolve their risks of CIN3+ to 6.5% (ASC-US+ cytology) and < 2% (cytology-negative). [The triage process can effectively risk-stratify women for the presence of underlying CIN3+. This example effectively illustrates context sensitivity and how the risk stratification inherent in the triage process must ultimately consider the underlying burden of disease as well as the local acceptability of various levels of risk.]

Table S1 (Annex 1; web only; available from https://publications.iarc.fr/604) shows the post-test risks of CIN3+ in triage-positive women (PPV) and in triage-negative women (cNPV) for all six triage strategies in low-risk, intermediate-risk, and high-risk situations. The green shading indicates, as an example, the decision thresholds chosen for risk of CIN3+ at > 10% for referral and < 1% for return to routine screening. [It should be noted that each local programme should choose its own decision thresholds in the context of locally acceptable risks. More complex algorithms than those assessed here can be considered to fine-tune management, particularly in relation to the management of an intermediate-risk group who are hrHPV-positive but have a negative triage test result at the index test, for whom surveillance (i.e. two-time triage testing) is an option (Arbyn et al., 2020).]

4.4.8 Harms of HPV testing

The harms of HPV testing consist of the psychosocial impact of screening and of a positive HPV test result, and the physical and psychosocial harms of the sampling procedure and of diagnostic follow-up procedures and

Fig. 4.7 Pre-test–post-test probability plot, showing the risk of CIN3+ through the triage pathway applied to hrHPV-positive women, computed from pooled accuracy estimates applied in a given pre-test risk situation



Triage with HPV16/18 genotyping followed by colposcopy if HPV16/18-positive. Women who are positive only for other hrHPV types are further triaged with cytology and referred for colposcopy if ASC-US+.
The first triage is applied to a median-risk situation with a pre-test risk of 8% (see left vertical axis). Applying HPV16/18 genotyping stratifies the risk to 19.5% if HPV16/18-positive and to 2.8% if positive only for other hrHPV types. Applying cytology to women who are positive only for other hrHPV types stratifies the risk to 6.5% if ASC-US+ and to 1.3% if cytology is normal.
ASC-US+, atypical squamous cells of undetermined significance or worse; CI, confidence interval; CIN3+, cervical intraepithelial neoplasia grade 3 or worse; Cyto, cytology; hrHPV, high-risk human papillomavirus; VIA, visual inspection with acetic acid.
Created by the Working Group.

treatments. The psychosocial impact of a positive HPV test result is potentially greater than that of an abnormal cytology result, because HPV is sexually transmitted. Qualitative information about psychosocial harms collected by focus groups and in-depth interviews (Anhang et al., 2004; Kahn et al., 2005; McCaffery et al., 2006; Waller et al., 2007; Daley et al., 2010; O'Connor et al., 2014; Patel et al., 2018) has revealed that a positive HPV test result may cause anxiety and distress and may lead to concerns about the association between HPV and cervical cancer. It may also evoke feelings of stigma and shame and influence sexual relationships by leading to feelings of blame or guilt towards previous or current sexual partners.

The psychosocial impact of HPV testing in cervical screening programmes has been estimated by questionnaire surveys. These include studies that measured harms of HPV testing as a primary screening test (McCaffery et al., 2004; Kitchener et al., 2008; Hsu et al., 2018; Andreassen et al., 2019; McBride et al., 2020) and studies that measured harms of HPV testing in women with ASC-US (Maissi et al., 2004; McCaffery et al., 2010; Kwan et al., 2011; Wang et al., 2011; Garcés-Palacio et al., 2018). To understand what type of information should be included in HPV screening invitation letters, in leaflets, and on websites in order to minimize psychosocial harms, several studies have examined whether the psychological harms experienced are influenced by a woman's knowledge about HPV (Waller et al., 2007; Papa et al., 2009; Burger et al., 2014; Markovic-Denic et al., 2018; Patel et al., 2018).

The harms associated with collection of samples may be different for clinician collection and sample collection at home using a self-sampling device. The experience with self-sampling has been assessed in questionnaire surveys (Nelson et al., 2017) containing items on the preference for self-sampling compared with clinician collection, and sometimes also items on the physical and/or psychosocial harms of the collection procedure.

Finally, the magnitude of the harms of HPV testing, diagnostic workup, and treatment of high-grade lesions in cervical screening can be represented by the numbers of screen-positive women, referrals for colposcopy, and treatments, and may be higher for HPV-based screening than for VIA or cytology-based screening because of the relatively high HPV test positivity rate in screening (Arbyn et al., 2012). The proportions of screen-positive women, referrals for colposcopy, and treatments have been reported in meta-analyses of diagnostic HPV screening studies, RCTs, and implementation studies of HPV screening. The magnitude of diagnostic and treatment harms of HPV DNA-based programmes compared with cytology-based and VIA-based programmes was presented in Sections 4.4.2 and 4.4.3, respectively.

(a) Psychosocial harms of HPV testing as a primary screening test

The first study on the psychosocial impact of HPV testing as a primary test in cervical screening was conducted in the United Kingdom in 271 women (mean age, 32 years) who received HPV testing and cytology testing (McCaffery et al., 2004). Anxiety was measured by the short form of the STAI-6 (Marteau & Bekker, 1992) and distress by the Cervical Screening Questionnaire (CSQ; Wardle et al., 1995), and results were collected within 1 month. Among women with normal cytology, anxiety and distress were higher in HPV-positive women than in HPV-negative women. A similar pattern was observed in women with abnormal or unsatisfactory cytology, but the variability of the estimates was high because the stratum size was only 40 women. In addition, more HPV-positive women than HPV-negative women felt worse about their current partner and about previous and future partners, and this effect was similar for women

with normal cytology and those with abnormal or unsatisfactory cytology.

Psychosocial outcomes in women with normal cytology were also measured in a substudy of the ARTISTIC trial (Kitchener et al., 2008, 2009a), a population-based randomized screening trial in the United Kingdom. Women with normal or mildly abnormal cytology recruited in the ARTISTIC trial were randomized either to cytology with revealed HPV testing or to cytology with concealed HPV testing. The women in the HPV-revealed arm received the results of their HPV test with their baseline cytology result; the women in the HPV-concealed arm were informed of only the cytology result. Anxiety, distress, and sexual satisfaction were assessed in 705 participants after about 2 weeks. Anxiety was measured by the STAI-6, distress was measured by the GHQ (Bridges & Goldberg, 1986), and sexual satisfaction was measured by the Sexual Rating Scale (Garratt et al., 1995). When the analysis was restricted to women who were aware of the HPV test result (the revealed arm) and who were cytology-negative, higher levels of anxiety and distress were reported in women who were HPV-positive than in women who were HPV-negative (41% vs 29%; OR, 1.70; 95% CI, 1.33–2.17). However, there was no evidence of a higher level of anxiety or distress in the revealed arm compared with the concealed arm (OR, 0.99; 95% CI, 0.81–1.21). A significant 7% difference on the Sexual Rating Scale was observed in HPV-positive women with normal cytology compared with the group of women with normal cytology and no revealed HPV test result.

A randomized implementation study of primary HPV screening versus cytology screening in Norway measured anxiety and depression by means of the Patient Health Questionnaire-4 (PHQ-4) (Kroenke et al., 2009) in 1007 screened women (Andreassen et al., 2019) randomized to either HPV testing every 5 years (followed by cytology if HPV-positive) or cytology testing every 3 years (followed by HPV testing if low-grade cytology was detected). Compared with women who were screened with cytology, women screened with an HPV test were not more likely to have mild, moderate, or severe anxiety and depression scores. Moreover, no differences in mean anxiety and depression levels were found when comparing HPV-positive women with normal cytology from the HPV screening group with women with normal cytology from the cytology group. [A possible explanation for the absence of an effect on psychosocial outcomes in the study in Norway is that women answered the questionnaire 4 months to 2 years after having received their last screening result, and elevations in anxiety and depression levels may have been temporary and levels may already have returned to normal. There was also considerable variation among participants in anxiety and depression levels, with some participants showing moderate or severe anxiety and depression levels.]

An inventory of the psychosocial harms in primary HPV screening implemented in a middle-income setting was conducted by Arrossi et al. (2020). In 163 HPV-positive women participating in the regional primary HPV screening programme in Jujuy, Argentina, psychosocial impact was measured by means of the Psycho-Estampa Scale, which was designed and validated for use in Latin American women. The Psycho-Estampa Scale consists of five domains: (i) an emotional domain, related to feelings about having a sexually transmitted infection; (ii) a sexuality domain, related to attitude and practice in sexual relationships; (iii) an uncertainty of information domain; (iv) a domain pertaining to the impact on family members; and (v) a worries domain, covering worries about HPV, cancer, and treatment. In the study population, the mean levels were highest for worries about HPV, cancer, and treatment but were also elevated for the other domains. The scores were higher in women with abnormal cytology triage than in women with normal cytology.

A systematic review of 25 studies on the effect of a positive HPV test on psychosexual outcomes (Bennett et al., 2019) considered overall psychosexual impact, sexual satisfaction and pleasure, frequency of sex, interest in sex, and feelings about partners and relationships. The studies included were very heterogeneous, which made it difficult to draw conclusions about the psychosexual impact of HPV testing, but in general women were concerned about transmitting HPV to a partner and about where the infection came from.

The longitudinal pattern of psychosocial outcomes was studied in England in a questionnaire survey in 1127 women aged 24–65 years who were screened at one of the primary HPV screening pilot centres; the study included a control group with negative cytology who were not tested for HPV (McBride et al., 2020). Elevated anxiety (STAI-6) and distress (GHQ) scores were recorded in HPV-positive women compared with women with negative cytology in the first 3 months after the test result had been received. However, after 12 months, anxiety and distress levels had returned to normal levels, irrespective of the HPV test result at 12 months. With respect to disease-related concerns, a positive HPV test result at baseline and at 12 months contributed to worry about cancer, and HPV clearance at 12 months contributed to reassurance. [The observation that a positive HPV test result at 12 months did not lead to an increase in the mean levels of anxiety and distress but was associated with worry about cancer suggests that although a positive HPV test result gives rise to disease-related concern initially, it is not disruptive of daily functioning when repeated.]

The observation that distress levels decrease over time was confirmed in a smaller study of 70 HPV-positive women in Taiwan, China, who were followed up until 12 months after a positive HPV test result (Hsu et al., 2018).

(b) Psychosocial harms of HPV testing as triage after an abnormal cytology result

One of the first studies that evaluated the psychosocial harms of HPV testing in women with an abnormal cytology result was a pilot study embedded in routine cytology screening in England, which recruited 1376 women with a normal or BMD cytology result (ASC-US/LSIL); 867 of the women with ASC-US/LSIL also had an HPV test (Maissi et al., 2004). The 536 women with a positive HPV test result were compared with the 331 women with a negative HPV test result and the 509 women who were not tested for HPV. Women with a positive HPV test result had the highest level of anxiety as measured by the STAI-6, the highest level of distress as measured by the GHQ, and the largest concern about the test result compared with the other groups. Women with an abnormal cytology result, whether tested for HPV or not, were less likely to know what their results meant compared with women with a normal cytology result; 26% of women with a positive HPV test result stated that they did not know what this meant for their health. Levels of anxiety, distress, and concern were similar in women with a negative HPV test result and in women who were not tested for HPV. [Because the study was cross-sectional, it did not provide information about the duration of elevated levels of anxiety and distress.] After a 6-month follow-up assessment (Maissi et al., 2005), mean levels of anxiety and distress were lower and did not differ between the three groups. The level of concern about a positive HPV test result was still elevated after 6 months compared with the level of concern after a negative HPV test result or no HPV test, but the level of concern had decreased from the baseline level. Worries about sexual health were measured for the first time after 6 months, and they were also higher in the group with a positive HPV test result.

An association between psychosocial harms and HPV testing does not necessarily imply that HPV triage has a negative effect on psychosocial outcomes in women with ASC-US. For example (as mentioned above) in the ARTISTIC trial, HPV-positive and HPV-negative women had different levels of psychosocial outcomes, but there were no significant differences in mean levels between the cytology and HPV randomization arms. To address this for women with ASC-US, in a pragmatic, randomized screening study in Australia of 314 women with an ASC-US test result, women were randomized to HPV testing, repeat cytology testing after 6 months, or an informed choice of either test supported by a decision tool (McCaffery et al., 2010). In the informed-choice arm, 61 (64%) women chose HPV testing and 35 (36%) chose repeat cytology testing. Psychosocial outcomes were measured after 2 weeks and after 3, 6, and 12 months. After 2 weeks, no mean effect of HPV testing was observed on anxiety as measured by the STAI-6 or on distress as measured by the CSQ (Wardle et al., 1995), although HPV testing was associated with 57% of women having intrusive thoughts in the HPV testing arm, compared with 32% in the repeat cytology testing arm and 43% in the informed-choice arm. However, after 1 year, most of the women in the HPV testing arm did not report residual intrusive thoughts, and distress was highest in the repeat cytology testing arm.

The temporary nature of anxiety, as observed in the studies in England and Australia described above, was confirmed in a study of 299 ethnic Chinese women in Hong Kong SAR with an ASC-US test result who received adjunct HPV testing (Kwan et al., 2011). Baseline differences in the mean level of anxiety (STAI-6) between HPV-negative and HPV-positive women had disappeared after 6 months. The effect of HPV testing on the HPV Impact Profile (HIP) was also examined. The HIP scale is a combined, multidimensional scale (Mast et al., 2009) with seven dimensions: worries and concerns, emotional impact, sexual impact, self-image, partner issues and transmission, interactions with physicians, and health control and impact on daily living. HIP scores were different for HPV-positive and HPV-negative women at baseline and at 6 months, although the differences were smaller at 6 months.

A hospital-based survey in China in 2605 women who had visited the hospital in the previous 3 months (Wang et al., 2011) confirmed that HIP scores were elevated in women with an HPV-positive ASC-US test result compared with women with an HPV-negative ASC-US test result or women with normal cytology. A pragmatic trial in Colombia compared psychosocial outcomes in 675 women (Garcés-Palacio et al., 2020) randomized to repeat cytology testing, HPV testing, or colposcopy after an ASC-US test result. The study found that anxiety measured by a long-form 20-item version of the Spielberger anxiety scale (STAI-20) and the HIP was higher in HPV-positive women than in HPV-negative women at 2 months, but that the differences in mean levels had disappeared after 1 year. There were no significant differences between the different randomization groups.

A strength of the randomized trials in Australia (McCaffery et al., 2010) and Colombia (Garcés-Palacio et al., 2020) is that the direct causal effect of HPV testing on psychosocial harms in the screening population is measured. This causal effect of learning about the HPV test result on psychosocial outcomes cannot be concluded from a comparison of psychosocial outcomes in HPV-positive and HPV-negative women, because HPV-positive women may have different levels of harms than HPV-negative women before the HPV test result is revealed. This conjecture was examined by a study in 2842 women in the United Kingdom (Johnson et al., 2011) participating in the TOMBOLA trial (Cotton et al., 2006). Psychosocial outcomes were measured before the HPV test result was

revealed. Anxiety was measured by the HADS (Zigmond & Snaith, 1983). In White women, there were no baseline differences in anxiety and cancer worries, but in non-White women, anxiety was lower in HPV-positive women than in HPV-negative women. In non-smokers, cancer worry was more common in HPV-positive women than in HPV-negative women; the opposite association was observed in ex-smokers.

[This suggests that the effect on psychosocial outcomes of knowing the HPV test result may be somewhat confounded by baseline differences between HPV-positive and HPV-negative women.]

(c) Psychosocial harms and knowledge about HPV

Mass education about HPV can prevent anxiety and psychological distress associated with HPV testing (Anhang et al., 2004). Focus group interviews (Anhang et al., 2005) identified that women desire detailed information about HPV, including susceptibility, risk of cervical cancer, and the effect of preventive interventions on this risk. The studies described here aimed to estimate the association between knowledge of HPV and psychosocial harms.

Waller et al. (2007) conducted a web-based survey in the United Kingdom in 811 female students. The participants were asked to imagine that they had had a positive HPV test result, and the study assessed the impact of their knowledge that HPV is sexually transmitted and about the high prevalence of HPV infection on stigma, shame, and anxiety by withholding pieces of information from some participants. Knowledge of the high prevalence was associated with lower levels of stigma, shame, and anxiety, whereas knowledge that HPV is sexually transmitted was associated with higher levels of stigma and shame but not anxiety. Women who knew that HPV is sexually transmitted but not that it is highly prevalent had the highest scores for stigma and shame.

The findings of this study were supported by a structured interview study in 46 women in the United Kingdom, which indicated that lack of knowledge enhances anxiety after a positive HPV test result (Patel et al., 2018), and a study of 324 women in Serbia with an abnormal cytology result (Markovic-Denic et al., 2018), which found that awareness of a positive HPV test result increases anxiety and perceived risk of cancer and concern, but that knowledge about HPV decreased anxiety and concern. Slightly different results were obtained by a small educational intervention study in the USA in 50 women aged 30 years and older (Papa et al., 2009), which indicated that education may not alleviate the concern about developing cancer, and a randomized web-based survey in 3540 women in Norway (Burger et al., 2014), which indicated that a switch to HPV screening does not increase anxiety, irrespective of whether additional information about HPV is provided.

[The study outcomes suggest that awareness that HPV is sexually transmitted increases levels of anxiety, stigma, and shame, but that low levels can be retained by creating awareness of the high prevalence of HPV. Implementation of HPV testing should be accompanied by a well-designed education and communication strategy to explain what a positive HPV test result means.]

(d) Diagnostic harms of HPV testing as triage after an ASC-US or LSIL test result

The magnitude of the diagnostic harms of HPV testing as triage is indicated by the clinical specificity for the absence of CIN2+ and the number of referrals for colposcopy. Pooled estimates were calculated in a meta-analysis of 39 studies in women with ASC-US and 24 studies in women with LSIL in whom HPV triage was conducted by HC2 testing; the women subsequently underwent colposcopy and colposcopy-directed biopsies for histological verification (Arbyn et al., 2012, 2013a). The pooled specificity of HPV triage testing after an ASC-US result

for detection of CIN2+ was 58.3% (95% CI, 53.6–62.9%). There was considerable variation across the studies, with specificities ranging from 27% to 79%. The pooled specificity of HPV triage testing for the management of LSIL for detection of CIN2+ was only 27.8% (95% CI, 23.8–32.1%) and varied from 16% to 58% across studies. The proportion of referrals for colposcopy was 48.2% (95% CI, 43.7–52.6%) for ASC-US and 76.9% (95% CI, 73.5–80.2%) for LSIL.

Three well-documented studies in the meta-analyses that were large enough to enable comparison of different age cohorts were the Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS) trial (Sherman et al., 2002), the NTCC trial (Ronco et al., 2007b), and the KPNC cohort (Castle et al., 2010). In women with ASC-US, the proportions of colposcopy referrals with HPV triage were 54% in the ALTS trial, 30% in the NTCC trial, and 35% in the KPNC cohort. In women with LSIL, the proportions of colposcopy referrals with HPV triage were 85% in the ALTS trial, 55% in the NTCC trial, and 84% in the KPNC cohort. In all three studies, the proportions of colposcopy referrals with HPV triage were dependent on age. In the ALTS trial, the proportion of women referred in the ASC-US subgroup decreased from 71% in women aged 18–22 years to 31% in women aged 29 years or older, whereas the referral proportion in the LSIL subgroup decreased only from 87% in women aged 18–22 years to 75% in women aged 29 years or older. In the NTCC trial, the referral proportions in the ASC-US subgroup were 46% in women aged 25–34 years and 25% in women aged 35–60 years, whereas the referral proportions in the LSIL subgroup were 72% in women aged 25–34 years and 41% in women aged 35–60 years. In the KPNC cohort, the referral proportions in the ASC-US subgroup decreased from 52% in women aged 30–34 years to 28% in women aged 60–64 years, and the referral proportions in the LSIL subgroup decreased from 89% in women aged 30–34 years to 74% in women aged 60–64 years.

(e) Psychosocial and physical harms of self-collection versus clinician collection

HPV testing can be performed on a self-collected sample, and this may decrease the physical and psychosocial harms of the sample collection process. Several studies have collected information about the impact of the sample collection method on the acceptability and harms of HPV testing. A systematic review of 20 studies that assessed the acceptability of self-sampling, preferences, and experience with self-sampling (Huynh et al., 2010) indicated that discomfort and pain were not experienced in general. Most women in the studies also had a positive attitude towards self-sampling as a part of future screening. A concern observed in multiple studies was that women were unsure whether they had followed the testing procedure correctly and had greater confidence in the accuracy of the clinician collection. The preference for self-sampling was also observed in a larger systematic review and meta-analysis of 37 studies published in 1986–2014 that included more than 18 000 women in North America, South America, Europe, Africa, and Asia (Nelson et al., 2017). Most of the studies were in countries in North America, South America, and Europe; six studies were in Asian countries, and five studies were in African countries. Nine studies involved self-sampling at home. The pooled estimate of women reporting a preference for self-collection over clinician collection was 59% (95% CI, 48–69%). Reasons for preferring self-collection were that it is easy to use and that it is private, not embarrassing, convenient, and comfortable. Some women reported that they disliked self-collection because it was painful or physically uncomfortable, because it led to anxiety, or because of uncertainty about whether the sampling was done correctly. Some women indicated that they did not like touching themselves. One study in women in India,

Nicaragua, and Uganda also reported that most women surveyed (78%) preferred self-sampling; 75% reported that it was easy, although 52% were initially concerned about hurting themselves and 24% were worried about not getting a good sample. The acceptability of self-sampling was higher when providers prepared the women through education, when providers allowed women to examine the collection brush, and when providers were present during the self-collection process (Bansil et al., 2014).

Since the two systematic reviews were conducted, several studies have been published in which women invited for HPV screening were asked about their experiences and/or harms of self-sampling. Most of those studies were pilot implementation studies evaluating home-based self-sampling, sometimes with the involvement of a community health worker. An overview of recent studies is given here. A study of home-based HPV self-sampling in 746 non-responders to the screening programme in Australia randomized women to self-collection for HPV testing or a repeat invitation letter for a cervical cytology test at the clinic (Sultana et al., 2015). More than 90% of the women considered self-collection to be easier, more convenient, less embarrassing, and less uncomfortable; however, similar to studies in the meta-analyses, most women were unsure about the reliability of the HPV self-sampling test result. Most women (88%) preferred self-sampling at home because it was simple and did not require an appointment at the clinician's office. Similar findings were reported in a study of home-based self-sampling with involvement of a community health worker in 200 underscreened Aboriginal women in rural and remote communities in Australia, more than 90% of whom indicated that they were highly satisfied with the HPV self-sampling kit and the process involved (Dutton et al., 2020). Two large studies in Latin America – a study in 2616 women in Argentina invited for regular screening (Arrossi et al., 2016) and a study in 1867 underscreened women in El Salvador (Maza et al., 2018) – assessed the attitude towards home-based self-sampling, both with involvement of a community health worker. Both studies reported that saving time was an additional reason to prefer self-sampling, in addition to the reasons that self-sampling is easy to perform and more comfortable and less embarrassing than clinician sampling. Maza et al. (2018) reported that feeling empowered was a reason for choosing self-sampling. Arrossi et al. (2016) reported, based on 433 women who chose clinician sampling instead of self-sampling, that the main reasons for not choosing self-sampling were trust in the clinician and the woman's fear of hurting herself. Another large self-sampling study included about 13 000 women in rural regions in Greece, who were recruited through a nationwide network of midwives (Chatzistamatiou et al., 2020). Women conducted self-sampling at home or at a general practitioner (GP) clinic and indicated minimal pain or discomfort and preference for self-collection when the test result is reliable. Testing at home was also preferred to self-sampling at a GP clinic. Positive experience of home-based self-sampling was also reported in other, smaller studies, including in women in rural Canada (Duke et al., 2015), Kenya (Oketch et al., 2019), Nigeria (Modibbo et al., 2017), and the United Republic of Tanzania (Bakiewicz et al., 2020), and in women in Japan with limited experience of tampon use (Hanley et al., 2016).

The role of home-based self-sampling in programmatic, regular screening is currently being discussed in several countries. In two recent studies, in the Netherlands (Polman et al., 2019c) and Sweden (Hermansson et al., 2020), HPV self-sampling was evaluated as a primary instrument in the setting of HPV-based screening without the use of an additional test for women with a negative HPV self-sampling result. In the study in the Netherlands (Polman et al., 2019c), experience was measured in routine screening in which women were randomized to HPV testing

on a self-collected versus clinician-collected sample. Responses were collected from 3835 women. Self-collection scored substantially lower on discomfort, pain, nervousness, and shame and higher on privacy compared with clinician collection. Trust in the test result was high with both self-collected and clinician-collected samples for HPV testing, irrespective of the HPV test result, although it was slightly higher for clinician sampling; 77% of the women reported that they preferred self-sampling for future screening. In the study in Sweden (Hermansson et al., 2020), in 868 women aged 60 years or older who had a positive HPV self-sampling result, 59% reported a preference for self-sampling versus 17% for clinician sampling. The main reasons for preferring self-sampling were that it is easy to perform and less embarrassing and less time-consuming than clinician sampling.

Information from non-responders and from clinicians can help to gain further insights into attitudes towards self-sampling. A study in underscreened women in the USA (Malone et al., 2020) compared attitudes in self-sampling kit returners (116 of 272 women invited) and non-returners (119 of 1083 women invited) and found no difference in attitude towards screening. The most common reason for non-return was low confidence in the woman's ability to correctly use the kit (Malone et al., 2020). In both groups, trust in the preventive effect of HPV screening against cancer was low. A randomized trial of HPV self-sampling in women in the USA that assessed attitudes in screened women and in clinicians (Mao et al., 2017) indicated that both screened women and clinicians expressed concerns about trust in the self-sampling test and valued the opportunity to discuss other health concerns with the clinician at the time of sampling.

Several individual studies compared attitudes and experiences with multiple sampling devices. In a study in non-responders in the Netherlands, the experiences of almost 10 000 women, to whom either a brush or lavage was offered, were compared (Bosgraaf et al., 2014). The experience of using the devices did not differ with respect to shame, feeling at ease, stress, discomfort, and pain, with levels similar to those observed in earlier studies. In a similarly designed study in Finland (Karjalainen et al., 2016), low discomfort and pain levels were reported for both devices. In a study in the KwaZulu-Natal region of South Africa in young women aged 16–22 years attending rural high schools (Mbatha et al., 2017), a choice between home-based self-sampling with a swab or a brush and clinician sampling was offered to all women. Most women expressed a preference for self-sampling (56%) compared with clinician sampling (44%). Pain was reported less often for the swab than for the brush, and the swab was preferred to the brush by most women who favoured self-sampling. However, in a study in Norway in women with a positive clinician-based hrHPV test, in which home-based self-sampling with a swab and a brush was subsequently offered to all women (Leinonen et al., 2018), both the swab and the brush were rated very positively, but the brush was reported as slightly easier to use and more comfortable.

[This indicates that although the experience was in general very positive, the preferred self-sampling method may vary across populations.]

References

Agorastos T, Chatzistamatiou K, Katsamagkas T, Koliopoulos G, Daponte A, Constantinidis T, et al.; HERMES study group (2015). Primary screening for cervical cancer based on high-risk human papillomavirus (HPV) detection and HPV 16 and HPV 18 genotyping, in comparison to cytology. *PLoS One*. 10(3):e0119755. doi:10.1371/journal.pone.0119755 PMID:25793281

Agorastos T, Chatzistamatiou K, Moysiadis T, Kaufmann AM, Skenderi A, Lekka I, et al. (2017). Human papillomavirus E7 protein detection as a method of triage to colposcopy of HPV positive women, in comparison to genotyping and cytology. Final results of the PIPAVIR

study. *Int J Cancer*. 141(3):519–30. doi:10.1002/ijc.30761 PMID:28470689

Agorastos T, Dinas K, Lloveras B, de Sanjose S, Kornegay JR, Bonti H, et al. (2005). Human papillomavirus testing for primary screening in women at low risk of developing cervical cancer. The Greek experience. *Gynecol Oncol*. 96(3):714–20. doi:10.1016/j.ygyno.2004.11.042 PMID:15721416

Aitken CA, van Agt HME, Siebers AG, van Kemenade FJ, Niesters HGM, Melchers WJG, et al. (2019). Introduction of primary screening using high-risk HPV DNA detection in the Dutch cervical cancer screening programme: a population-based cohort study. *BMC Med*. 17(1):228. doi:10.1186/s12916-019-1460-0 PMID:31829241

Almonte M, Murillo R, Sánchez GI, González P, Ferrera A, Picconi MA, et al. (2020). Multicentric study of cervical cancer screening with human papillomavirus testing and assessment of triage methods in Latin America: the ESTAMPA screening study protocol. *BMJ Open*. 10(5):e035796. doi:10.1136/bmjopen-2019-035796 PMID:32448795

Andreassen T, Hansen BT, Engesæter B, Hashim D, Støer NC, Tropé A, et al. (2019). Psychological effect of cervical cancer screening when changing primary screening method from cytology to high-risk human papilloma virus testing. *Int J Cancer*. 145(1):29–39. doi:10.1002/ijc.32067 PMID:30549273

Anhang R, Nelson JA, Telerant R, Chiasson MA, Wright TC Jr (2005). Acceptability of self-collection of specimens for HPV DNA testing in an urban population. *J Womens Health (Larchmt)*. 14(8):721–8. doi:10.1089/jwh.2005.14.721 PMID:16232104

Anhang R, Wright TC Jr, Smock L, Goldie SJ (2004). Women's desired information about human papillomavirus. *Cancer*. 100(2):315–20. doi:10.1002/cncr.20007 PMID:14716766

Arbyn M, Hillemanns P (2018). HPV assays validated for primary cervical cancer screening. *HPV World*. 55:6–9. Available from: https://www.hpvworld.com/articles/hpv-assays-validated-for-primary-cervical-cancer-screening/.

Arbyn M, Rezhake R, Yuill S, Canfell K (2020). Triage of HPV-positive women in Norway using cytology, HPV16/18 genotyping and HPV persistence. *Br J Cancer*. 122(11):1577–9. doi:10.1038/s41416-020-0787-9 PMID:32242099

Arbyn M, Roelens J, Cuschieri K, Cuzick J, Szarewski A, Ratnam S, et al. (2013b). The APTIMA HPV assay versus the Hybrid Capture 2 test in triage of women with ASC-US or LSIL cervical cytology: a meta-analysis of the diagnostic accuracy. *Int J Cancer*. 132(1):101–8. doi:10.1002/ijc.27636 PMID:22610699

Arbyn M, Roelens J, Simoens C, Buntinx F, Paraskevaidis E, Martin-Hirsch PPL, et al. (2013a). Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions. *Cochrane Database Syst Rev*. 2013(3):CD008054. doi:10.1002/14651858.CD008054.pub2 PMID:23543559

Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G, et al. (2012). Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. [Erratum in: Vaccine. 2013 31(52):6266] *Vaccine*. 30(Suppl 5):F88–99. doi:10.1016/j.vaccine.2012.06.095 PMID:23199969

Arbyn M, Sankaranarayanan R, Muwonge R, Keita N, Dolo A, Mbalawa CG, et al. (2008). Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *Int J Cancer*. 123(1):153–60. doi:10.1002/ijc.23489 PMID:18404671

Arbyn M, Smith SB, Temin S, Sultana F, Castle P; Collaboration on Self-Sampling and HPV Testing (2018). Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. *BMJ*. 363:k4823. doi:10.1136/bmj.k4823 PMID:30518635

Arbyn M, Snijders PJF, Meijer CJLM, Berkhof J, Cuschieri K, Kocjan BJ, et al. (2015). Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clin Microbiol Infect*. 21(9):817–26. doi:10.1016/j.cmi.2015.04.015 PMID:25936581

Arbyn M, Verdoodt F, Snijders PJF, Verhoef VMJ, Suonio E, Dillner L, et al. (2014). Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol*. 15(2):172–83. doi:10.1016/S1470-2045(13)70570-9 PMID:24433684

Arbyn M, Xu L, Verdoodt F, Cuzick J, Szarewski A, Belinson JL, et al. (2017). Genotyping for human papillomavirus types 16 and 18 in women with minor cervical lesions: a systematic review and meta-analysis. *Ann Intern Med*. 166(2):118–27.doi:10.7326/M15-2735 PMID:27842420

Arrossi S, Almonte M, Herrero R, Gago J, Sánchez Antelo V, Szwarc L, et al. (2020). Psycho-social impact of positive human papillomavirus testing in Jujuy, Argentina results from the Psycho-Estampa study. *Prev Med Rep*. 18:101070. doi:10.1016/j.pmedr.2020.101070 PMID:32257775

Arrossi S, Paolino M, Laudi R, Gago J, Campanera A, Marín O, et al. (2019). Programmatic human papillomavirus testing in cervical cancer prevention in the Jujuy Demonstration Project in Argentina: a population-based, before-and-after retrospective cohort study. *Lancet Glob Health*. 7(6):e772–83. doi:10.1016/S2214-109X(19)30048-8 PMID:31097279

Arrossi S, Ramos S, Straw C, Thouyaret L, Orellana L (2016). HPV testing: a mixed-method approach to understand why women prefer self-collection in a middle-income country. *BMC Public Health*. 16(1):832. doi:10.1186/s12889-016-3474-2 PMID:27538390

Asthana S, Labani S (2015). Adjunct screening of cervical or vaginal samples using careHPV testing with Pap and aided visual inspection for detecting high-grade cervical intraepithelial neoplasia. *Cancer Epidemiol.* 39(1):104–8. doi:10.1016/j.canep.2014.11.006 PMID: 25548076

Bakiewicz A, Rasch V, Mwaiselage J, Linde DS (2020). "The best thing is that you are doing it for yourself" – perspectives on acceptability and feasibility of HPV self-sampling among cervical cancer screening clients in Tanzania: a qualitative pilot study. *BMC Womens Health.* 20(1):65. doi:10.1186/s12905-020-00917-7 PMID:32234028

Bansil P, Wittet S, Lim JL, Winkler JL, Paul P, Jeronimo J (2014). Acceptability of self-collection sampling for HPV-DNA testing in low-resource settings: a mixed methods approach. *BMC Public Health.* 14(1):596. doi:10.1186/1471-2458-14-596 PMID:24927941

Basu P, Mittal S, Banerjee D, Singh P, Panda C, Dutta S, et al. (2015). Diagnostic accuracy of VIA and HPV detection as primary and sequential screening tests in a cervical cancer screening demonstration project in India. *Int J Cancer.* 137(4):859–67. doi:10.1002/ijc.29458 PMID:25631198

Belinson J, Qiao YL, Pretorius R, Zhang WH, Elson P, Li L, et al. (2001). Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. [Erratum in: Gynecol Oncol. 2002 84(2):355] *Gynecol Oncol.* 83(2):439–44. doi:10.1006/gyno.2001.6370 PMID:11606114

Belinson JL, Qiao YL, Pretorius RG, Zhang WH, Rong SD, Huang MN, et al. (2003). 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. 13(6):819–26. doi:10.1136/ijgc-00009577-200311000-00013 PMID:14675319

Belinson SE, Wulan N, Li R, Zhang W, Rong X, Zhu Y, et al. (2010). SNIPER: a novel assay for human papillomavirus testing among women in Guizhou, China. *Int J Gynecol Cancer.* 20(6):1006–10. doi:10.1111/IGC.0b013e3181e73092 PMID:20683409

Bennett KF, Waller J, Ryan M, Bailey JV, Marlow LAV (2019). The psychosexual impact of testing positive for high-risk cervical human papillomavirus (HPV): a systematic review. *Psychooncology.* 28(10):1959–70. doi:10.1002/pon.5198 PMID:31411787

Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology.* 401(1):70–9. doi:10.1016/j.virol.2010.02.002 PMID:20206957

Bian ML, Cheng JY, Ma L, Cong X, Liu J, Chen Y, et al. (2013). Evaluation of the detection of 14 high-risk human papillomaviruses with HPV 16 and HPV 18 genotyping for cervical cancer screening. *Exp Ther Med.* 6(5):1332–6. doi:10.3892/etm.2013.1309 PMID:24223668

Bigras G, de Marval F (2005). The probability for a Pap test to be abnormal is directly proportional to HPV viral load: results from a Swiss study comparing HPV testing and liquid-based cytology to detect cervical cancer precursors in 13,842 women. *Br J Cancer.* 93(5):575–81. doi:10.1038/sj.bjc.6602728 PMID:16136031

Blatt AJ, Kennedy R, Luff RD, Austin RM, Rabin DS (2015). Comparison of cervical cancer screening results among 256,648 women in multiple clinical practices. *Cancer Cytopathol.* 123(5):282–8. doi:10.1002/cncy.21544 PMID:25864682

Bobdey S, Balasubramanium G, Kumar A, Jain A (2015). Cancer screening: should cancer screening be essential component of primary health care in developing countries? *Int J Prev Med.* 6:56. PMID:26236443

Bobdey S, Sathwara J, Jain A, Balasubramaniam G (2016). Burden of cervical cancer and role of screening in India. *Indian J Med Paediatr Oncol.* 37(4):278–85. doi:10.4103/0971-5851.195751 PMID:28144096

Bosgraaf RP, Ketelaars PJW, Verhoef VMJ, Massuger LFAG, Meijer CJLM, Melchers WJG, et al. (2014). Reasons for non-attendance to cervical screening and preferences for HPV self-sampling in Dutch women. *Prev Med.* 64:108–13. doi:10.1016/j.ypmed.2014.04.011 PMID:24736093

Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al.; WHO International Agency for Research on Cancer Monograph Working Group (2009). A review of human carcinogens – Part B: biological agents. *Lancet Oncol.* 10(4):321–2. doi:10.1016/S1470-2045(09)70096-8 PMID:19350698

Bridges KW, Goldberg DP (1986). The validation of the GHQ-28 and the use of the MMSE in neurological in-patients. *Br J Psychiatry.* 148(5):548–53. doi:10.1192/bjp.148.5.548 PMID:3779225

Bulkmans NWJ, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJP, Bulk S, et al. (2007). 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.* 370(9601):1764–72. doi:10.1016/S0140-6736(07)61450-0 PMID:17919718

Bulkmans NWJ, Rozendaal L, Snijders PJF, Voorhorst FJ, Boeke AJP, Zandwijken GRJ, et al. (2004). 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.* 110(1):94–101. doi:10.1002/ijc.20076 PMID:15054873

Burger EA, Nygård M, Gyrd-Hansen D, Moger TA, Kristiansen IS (2014). Does the primary screening test influence women's anxiety and intention to screen for cervical cancer? A randomized survey of Norwegian women. *BMC Public Health*. 14(1):360. doi:10.1186/1471-2458-14-360 PMID:24735469

Campos NG, Alfaro K, Maza M, Sy S, Melendez M, Masch R, et al. (2020). The cost-effectiveness of human papillomavirus self-collection among cervical cancer screening non-attenders in El Salvador. *Prev Med*. 131:105931. doi:10.1016/j.ypmed.2019.105931 PMID:31765712

Campos NG, Tsu V, Jeronimo J, Njama-Meya D, Mvundura M, Kim JJ (2017). Cost-effectiveness of an HPV self-collection campaign in Uganda: comparing models for delivery of cervical cancer screening in a low-income setting. *Health Policy Plan*. 32(7):956–68. doi:10.1093/heapol/czw182 PMID:28369405

Canfell K, Caruana M, Gebski V, Darlington-Brown J, Heley S, Brotherton J, et al. (2017). Cervical screening with primary HPV testing or cytology in a population of women in which those aged 33 years or younger had previously been offered HPV vaccination: results of the Compass pilot randomised trial. *PLoS Med*. 14(9):e1002388. doi:10.1371/journal.pmed.1002388 PMID:28926579

Canfell K, Saville M, Caruana M, Gebski V, Darlington-Brown J, Brotherton J, et al. (2018). Protocol for Compass: a randomised controlled trial of primary HPV testing versus cytology screening for cervical cancer in HPV-unvaccinated and vaccinated women aged 25–69 years living in Australia. [Erratum in: BMJ Open. 2018 8(3):e016700corr1] *BMJ Open*. 8(1):e016700. doi:10.1136/bmjopen-2017-016700 PMID:29374658

Cárdenas-Turanzas M, Nogueras-Gonzalez GM, Scheurer ME, Adler-Storthz K, Benedet JL, Beck JR, et al. (2008). The performance of human papillomavirus high-risk DNA testing in the screening and diagnostic settings. *Cancer Epidemiol Biomarkers Prev*. 17(10):2865–71. doi:10.1158/1055-9965.EPI-08-0137 PMID:18843032

Castle PE (2015). Comparison of cervical cancer screening results among 256,648 women in multiple clinical practices. *Cancer Cytopathol*. 123(9):566. doi:10.1002/cncy.21572 PMID:26179787

Castle PE, Fetterman B, Thomas Cox J, Shaber R, Poitras N, Lorey T, et al. (2010). The age-specific relationships of abnormal cytology and human papillomavirus DNA results to the risk of cervical precancer and cancer. *Obstet Gynecol*. 116(1):76–84. doi:10.1097/AOG.0b013e3181e3e719 PMID:20567171

Castle PE, Kinney WK, Xue X, Cheung LC, Gage JC, Poitras NE, et al. (2019). Role of screening history in clinical meaning and optimal management of positive cervical screening results. *J Natl Cancer Inst*. 111(8):820–7. doi:10.1093/jnci/djy192 PMID:30576462

Castle PE, Kinney WK, Xue X, Cheung LC, Gage JC, Zhao FH, et al. (2018). Effect of several negative rounds of human papillomavirus and cytology co-testing on safety against cervical cancer: an observational cohort study. *Ann Intern Med*. 168(1):20–9. doi:10.7326/M17-1609 PMID:29181509

Castle PE, Stoler MH, Wright TC Jr, Sharma A, Wright TL, Behrens CM (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(9):880–90. doi:10.1016/S1470-2045(11)70188-7 PMID:21865084

Chan KKL, Liu SS, Wei N, Ngu SF, Chu MMY, Tse KY, et al. (2020). Primary HPV testing with cytology versus cytology alone in cervical screening – a prospective randomized controlled trial with two rounds of screening in a Chinese population. *Int J Cancer*. 147(4):1152–62. doi:10.1002/ijc.32861 PMID:31922265

Chatzistamatiou K, Vrekoussis T, Tsertanidou A, Moysiadis T, Mouchtaropoulou E, Pasentsis K, et al. (2020). Acceptability of self-sampling for human papillomavirus-based cervical cancer screening. *J Womens Health (Larchmt)*. 29(11):1447–56. doi:10.1089/jwh.2019.8258 PMID:32757997

Chen C, Yang Z, Li Z, Li L (2012). Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis. *Int J Gynecol Cancer*. 22(6):908–21. doi:10.1097/IGC.0b013e318256e5e4 PMID:22672987

Cheung LC, Egemen D, Chen X, Katki HA, Demarco M, Wiser AL, et al. (2020). 2019 ASCCP risk-based management consensus guidelines: methods for risk estimation, recommended management, and validation. *J Low Genit Tract Dis*. 24(2):90–101. doi:10.1097/LGT.0000000000000528 PMID:32243306

Choi JW, Kim Y, Lee JH, Kim YS (2016). The clinical performance of primary HPV screening, primary HPV screening plus cytology cotesting, and cytology alone at a tertiary care hospital. *Cancer Cytopathol*. 124(2):144–52. doi:10.1002/cncy.21632 PMID:26457676

Clavel C, Masure M, Bory JP, Putaud I, Mangeonjean C, Lorenzato M, et al. (2001). Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: a study of 7932 women. *Br J Cancer*. 84(12):1616–23. doi:10.1054/bjoc.2001.1845 PMID:11401314

Coldman AJ, van Niekerk D, Krajden M, Smith LW, Cook D, Gondara L, et al. (2020). Disease detection at the 48-month exit round of the HPV FOCAL cervical cancer screening trial in women per-protocol eligible for routine screening. *Int J Cancer*. 146(7):1810–8. doi:10.1002/ijc.32524 PMID:31245842

Cook DA, Smith LW, Law J, Mei W, van Niekerk DJ, Ceballos K, et al. (2017). Aptima HPV Assay versus Hybrid Capture® 2 HPV test for primary cervical

cancer screening in the HPV FOCAL trial. *J Clin Virol*. 87:23–9. doi:10.1016/j.jcv.2016.12.004 PMID:27988420

Cook DA, Smith LW, Law JH, Mei W, Gondara L, van Niekerk DJ, et al. (2018). Comparative performance of human papillomavirus messenger RNA versus DNA screening tests at baseline and 48 months in the HPV FOCAL trial. *J Clin Virol*. 108:32–7. doi:10.1016/j.jcv.2018.09.004 PMID:30223252

Cotton SC, Sharp L, Little J, Duncan I, Alexander L, Cruickshank ME, et al.; TOMBOLA group (2006). Trial of management of borderline and other low-grade abnormal smears (TOMBOLA): trial design. *Contemp Clin Trials*. 27(5):449–71. doi:10.1016/j.cct.2006.04.001 PMID:16765101

Cuzick J, Szarewski A, Cubie H, Hulman G, Kitchener H, Luesley D, et al. (2003). Management of women who test positive for high-risk types of human papillomavirus: the HART study. *Lancet*. 362(9399):1871–6. doi:10.1016/S0140-6736(03)14955-0 PMID:14667741

Daley EM, Perrin KM, McDermott RJ, Vamos CA, Rayko HL, Packing-Ebuen JL, et al. (2010). The psychosocial burden of HPV: a mixed-method study of knowledge, attitudes and behaviors among HPV+ women. *J Health Psychol*. 15(2):279–90. doi:10.1177/1359105309351249 PMID:20207671

de Cremoux P, Coste J, Sastre-Garau X, Thioux M, Bouillac C, Labbé S, et al.; French Society of Clinical Cytology Study Group (2003). Efficiency of the Hybrid Capture 2 HPV DNA test in cervical cancer screening. A study by the French Society of Clinical Cytology. *Am J Clin Pathol*. 120(4):492–9. doi:10.1309/XFUCPP6M5XUA94B8 PMID:14560561

Del Mistro A, Giorgi Rossi P, Frayle H, Pasquale L, Campari C, Ronco G, et al. (2019). Five-year risk of CIN3 after short-term HPV-DNA negativity in cytology-negative women: a population-based cohort study. *BJOG*. 126(11):1365–71. doi:10.1111/1471-0528.15893 PMID:31356722

Demarco M, Egemen D, Raine-Bennett TR, Cheung LC, Befano B, Poitras NE, et al. (2020). A study of partial human papillomavirus genotyping in support of the 2019 ASCCP Risk-Based Management Consensus Guidelines. *J Low Genit Tract Dis*. 24(2):144–7. doi:10.1097/LGT.0000000000000530 PMID:32243309

Denny L, Kuhn L, De Souza M, Pollack AE, Dupree W, Wright TC Jr (2005). Screen-and-treat approaches for cervical cancer prevention in low-resource settings: a randomized controlled trial. *JAMA*. 294(17):2173–81. doi:10.1001/jama.294.17.2173 PMID:16264158

Denny L, Kuhn L, Hu CC, Tsai WY, Wright TC Jr (2010). Human papillomavirus-based cervical cancer prevention: long-term results of a randomized screening trial. *J Natl Cancer Inst*. 102(20):1557–67. doi:10.1093/jnci/djq342 PMID:20884893

Depuydt CE, Makar AP, Ruymbeke MJ, Benoy IH, Vereecken AJ, Bogers JJ (2011). BD-ProExC as adjunct molecular marker for improved detection of CIN2+ after HPV primary screening. *Cancer Epidemiol Biomarkers Prev*. 20(4):628–37. doi:10.1158/1055-9965.EPI-10-0818 PMID:21297040

Dijkstra MG, van Zummeren M, Rozendaal L, van Kemenade FJ, Helmerhorst TJM, Snijders PJF, et al. (2016). Safety of extending screening intervals beyond five years in cervical screening programmes with testing for high risk human papillomavirus: 14 year follow-up of population based randomised cohort in the Netherlands. *BMJ*. 355:i4924. doi:10.1136/bmj.i4924 PMID:27702796

Dillner J, Arbyn M, Unger E, Dillner L (2011). Monitoring of human papillomavirus vaccination. *Clin Exp Immunol*. 163(1):17–25. doi:10.1111/j.1365-2249.2010.04268.x PMID:21062269

Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. (2008). Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ*. 337:a1754. doi:10.1136/bmj.a1754 PMID:18852164

Duke P, Godwin M, Ratnam S, Dawson L, Fontaine D, Lear A, et al. (2015). Effect of vaginal self-sampling on cervical cancer screening rates: a community-based study in Newfoundland. *BMC Womens Health*. 15(1):47. doi:10.1186/s12905-015-0206-1 PMID:26060041

Dutton T, Marjoram J, Burgess S, Montgomery L, Vail A, Callan N, et al. (2020). Uptake and acceptability of human papillomavirus self-sampling in rural and remote Aboriginal communities: evaluation of a nurse-led community engagement model. *BMC Health Serv Res*. 20(1):398. doi:10.1186/s12913-020-05214-5 PMID:32393243

Dyer N, Young L, Ott S (2016). Artifacts in the data of Hu et al. *Nat Genet*. 48(1):2–4. doi:10.1038/ng.3392 PMID:26711106

Eklund C, Forslund O, Wallin KL, Dillner J (2014). Global improvement in genotyping of human papillomavirus DNA: the 2011 HPV LabNet International Proficiency Study. *J Clin Microbiol*. 52(2):449–59. doi:10.1128/JCM.02453-13 PMID:24478473

El-Zein M, Bouten S, Louvanto K, Gilbert L, Gotlieb W, Hemmings R, et al.; CASSIS Study Group (2018). Validation of a new HPV self-sampling device for cervical cancer screening: the Cervical and Self-Sample In Screening (CASSIS) study. *Gynecol Oncol*. 149(3):491–7. doi:10.1016/j.ygyno.2018.04.004 PMID:29678360

Elfgren K, Elfström KM, Naucler P, Arnheim-Dahlström L, Dillner J (2017). Management of women with human papillomavirus persistence: long-term follow-up of a randomized clinical trial. *Am J Obstet Gynecol*. 216(3):264.e1–7. doi:10.1016/j.ajog.2016.10.042 PMID:27825977

Elfström KM, Smelov V, Johansson AL, Eklund C, Nauclér P, Arnheim-Dahlström L, et al. (2014). Long term duration of protective effect for HPV negative women: follow-up of primary HPV screening randomised controlled trial. *BMJ*. 348:g130. doi:10.1136/bmj.g130 PMID:24435414

Ferreccio C, Barriga MI, Lagos M, Ibáñez C, Poggi H, González F, et al. (2013). Screening trial of human papillomavirus for early detection of cervical cancer in Santiago, Chile. *Int J Cancer*. 132(4):916–23. doi:10.1002/ijc.27662 PMID:22684726

Fokom-Domgue J, Combescure C, Fokom-Defo V, Tebeu PM, Vassilakos P, Kengne AP, et al. (2015). Performance of alternative strategies for primary cervical cancer screening in sub-Saharan Africa: systematic review and meta-analysis of diagnostic test accuracy studies. *BMJ*. 351:h3084. doi:10.1136/bmj.h3084 PMID:26142020

Forslund O, Elfström KM, Lamin H, Dillner J (2019). HPV-mRNA and HPV-DNA detection in samples taken up to seven years before severe dysplasia of cervix uteri. *Int J Cancer*. 144(5):1073–81. doi:10.1002/ijc.31819 PMID:30125346

Gage JC, Schiffman M, Katki HA, Castle PE, Fetterman B, Wentzensen N, et al. (2014). Reassurance against future risk of precancer and cancer conferred by a negative human papillomavirus test. *J Natl Cancer Inst*. 106(8):dju153. doi:10.1093/jnci/dju153 PMID:25038467

Garcés-Palacio IC, Ramos-Jaraba SM, Rubio-León DC (2018). Health beliefs associated with the follow-up of Pap smear abnormalities among low-income women in Medellín, Colombia. *J Cancer Educ*. 33(2):417–23. doi:10.1007/s13187-017-1172-0 PMID:28120138

Garcés-Palacio IC, Sanchez GI, Baena Zapata A, Córdoba Sánchez V, Urrea Cosme Y, Rodríguez Zabala D, et al. (2020). Psychosocial impact of inclusion of HPV test on the management of women with atypical squamous cells of undetermined significance: a study within a randomised pragmatic trial in a middle-income country. *Psychol Health*. 35(6):750–69. doi:10.1080/08870446.2019.1678749 PMID:31625399

Garratt AM, Torgerson DJ, Wyness J, Hall MH, Reid DM (1995). Measuring sexual functioning in premenopausal women. *Br J Obstet Gynaecol*. 102(4):311–6. doi:10.1111/j.1471-0528.1995.tb09138.x PMID:7612515

Gen-Probe (2011). APTIMA® HPV Assay. For in vitro diagnostic use. San Diego (CA), USA: Gen-Probe Inc. Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf10/P100042c.pdf.

Giorgi Rossi P, Ronco G, Dillner J, Elfström KM, Snijders PJF, Arbyn M, et al.; New Technologies for Cervical Cancer Group (2016). Why follow-back studies should be interpreted cautiously: the case of an HPV-negative cervical lesion. *Cancer Cytopathol*. 124(1):66–7. doi:10.1002/cncy.21622 PMID:26356132

Gravitt PE, Paul P, Katki HA, Vendantham H, Ramakrishna G, Sudula M, et al.; CATCH Study Team (2010). Effectiveness of VIA, Pap, and HPV DNA testing in a cervical cancer screening program in a peri-urban community in Andhra Pradesh, India. *PLoS One*. 5(10):e13711. doi:10.1371/journal.pone.0013711 PMID:21060889

Gustinucci D, Giorgi Rossi P, Cesarini E, Broccolini M, Bulletti S, Carlani A, et al. (2016). Use of cytology, E6/E7 mRNA, and p16^{INK4a}-Ki-67 to define the management of human papillomavirus (HPV)-positive women in cervical cancer screening. *Am J Clin Pathol*. 145(1):35–45. doi:10.1093/ajcp/aqv019 PMID:26712869

Haedicke J, Iftner T (2016). A review of the clinical performance of the Aptima HPV assay. *J Clin Virol*. 76(Suppl 1):S40–8. doi:10.1016/j.jcv.2015.10.027 PMID:26614686

Han L, Chang X, Song P, Gao L, Zhang Y, An L, et al. (2020). An on-going study of three different cervical cancer screening strategies based on primary healthcare facilities in Beijing China. *J Infect Public Health*. 13(4):577–83. doi:10.1016/j.jiph.2019.09.003 PMID:31564529

Hanley SJ, Fujita H, Yokoyama S, Kunisawa S, Tamakoshi A, Dong P, et al. (2016). HPV self-sampling in Japanese women: a feasibility study in a population with limited experience of tampon use. *J Med Screen*. 23(3):164–70. doi:10.1177/0969141315625702 PMID:26919866

Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC (2007). A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 8(2):239–51. doi:10.1093/biostatistics/kxl004 PMID:16698768

Harbord RM, Whiting P (2009). metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J*. 9(2):211–29. doi:10.1177/1536867X0900900203

HAS (2019). Evaluation de la recherche des papillomavirus humains (HPV) en dépistage primaire des lésions précancéreuses et cancéreuses du col de l'utérus et de la place du double immuno-marquage p16/Ki67. Saint-Denis La Plaine, France: Haute Autorité de Santé. [in French] Available from: https://www.has-sante.fr/upload/docs/application/pdf/2019-09/rapport_hpv_2019-09-09_14-20-44_251.pdf.

Herbert A, Arbyn M, Bergeron C (2008). Why CIN3 and CIN2 should be distinguished on histological reports. *Cytopathology*. 19(1):63–4. doi:10.1111/j.1365-2303.2006.00539.x PMID:18205632

Hermansson RS, Olovsson M, Gustavsson C, Lindström AK (2020). Elderly women's experiences of self-sampling for HPV testing. *BMC Cancer*. 20(1):473. doi:10.1186/s12885-020-06977-0 PMID:32456679

Holt HK, Zhang L, Zhao FH, Hu SY, Zhao XL, Zhang X, et al. (2017). Evaluation of multiple primary and combination screening strategies in postmenopausal women for detection of cervical cancer in China. *Int J Cancer*. 140(3):544–54. doi:10.1002/ijc.30468 PMID:27727464

Horn J, Denecke A, Luyten A, Rothe B, Reinecke-Lüthge A, Mikolajczyk R, et al. (2019). Reduction of cervical cancer incidence within a primary HPV screening pilot project (WOLPHSCREEN) in Wolfsburg, Germany. *Br J Cancer*. 120(10):1015–22. doi:10.1038/s41416-019-0453-2 PMID:30988395

Hovland S, Arbyn M, Lie AK, Ryd W, Borge B, Berle EJ, et al. (2010). A comprehensive evaluation of the accuracy of cervical pre-cancer detection methods in a high-risk area in East Congo. *Br J Cancer*. 102(6):957–65. doi:10.1038/sj.bjc.6605594 PMID:20197765

Hsu YY, Wang WM, Fetzer SJ, Cheng YM, Hsu KF (2018). Longitudinal psychosocial adjustment of women to human papillomavirus infection. *J Adv Nurs*. 74(11):2523–32. doi:10.1111/jan.13725 PMID:29845650

Hu Z, Zhu D, Wang W, Li W, Jia W, Zeng X, et al. (2015). Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. *Nat Genet*. 47(2):158–63. doi:10.1038/ng.3178 PMID:25581428

Huh WK, Ault KA, Chelmow D, Davey DD, Goulart RA, Garcia FA, et al. (2015). Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *Obstet Gynecol*. 125(2):330–7. doi:10.1097/AOG.0000000000000669 PMID:25569009

Huynh J, Howard M, Lytwyn A (2010). Self-collection for vaginal human papillomavirus testing: systematic review of studies asking women their perceptions. *J Low Genit Tract Dis*. 14(4):356–62. doi:10.1097/LGT.0b013e3181dc115b PMID:20885165

IARC (2005). Cervix cancer screening. *IARC Handb Cancer Prev*. 10:1–302. Available from: https://publications.iarc.fr/380.

IARC (2007). Human papillomaviruses. *IARC Monogr Eval Carcinog Risks Hum*. 90:1–670. Available from: https://publications.iarc.fr/108 PMID:18354839

IARC (2012). Biological agents. *IARC Monogr Eval Carcinog Risks Hum*. 100B:1–475. Available from: https://publications.iarc.fr/119 PMID:23189750

Iftner T, Becker S, Neis KJ, Castanon A, Iftner A, Holz B, et al. (2015). Head-to-head comparison of the RNA-based Aptima human papillomavirus (HPV) assay and the DNA-based Hybrid Capture 2 HPV test in a routine screening population of women aged 30 to 60 years in Germany. *J Clin Microbiol*. 53(8):2509–16. doi:10.1128/JCM.01013-15 PMID:26019212

Iftner T, Neis KJ, Castanon A, Landy R, Holz B, Woll-Herrmann A, et al. (2019). Longitudinal clinical performance of the RNA-based Aptima human papillomavirus (AHPV) assay in comparison to the DNA-based Hybrid Capture 2 HPV test in two consecutive screening rounds with a 6-year interval in Germany. *J Clin Microbiol*. 57(1):e01177–18. doi:10.1128/JCM.01177-18 PMID:30355760

Iftner T, Villa LL (2003). Chapter 12: human papillomavirus technologies. *J Natl Cancer Inst Monogr*. 2003(31):80–8. doi:10.1093/oxfordjournals.jncimonographs.a003487 PMID:12807950

Isidean SD, Mayrand MH, Ramanakumar AV, Gilbert L, Reid SL, Rodrigues I, et al.; CCCaST Study Group (2016). Human papillomavirus testing versus cytology in primary cervical cancer screening: end-of-study and extended follow-up results from the Canadian Cervical Cancer Screening Trial. *Int J Cancer*. 139(11):2456–66. doi:10.1002/ijc.30385 PMID:27538188

Isidean SD, Mayrand MH, Ramanakumar AV, Rodrigues I, Ferenczy A, Ratnam S, et al.; CCCaST Study Group (2017). Comparison of triage strategies for HPV-positive women: Canadian Cervical Cancer Screening Trial results. *Cancer Epidemiol Biomarkers Prev*. 26(6):923–9. doi:10.1158/1055-9965.EPI-16-0705 PMID:28096198

Islam JY, Mutua MM, Kabare E, Manguro G, Hudgens MG, Poole C, et al. (2020). High-risk human papillomavirus messenger RNA testing in wet and dry self-collected specimens for high-grade cervical lesion detection in Mombasa, Kenya. *Sex Transm Dis*. 47(7):464–72. doi:10.1097/OLQ.0000000000001167 PMID:32168174

Johnson CY, Sharp L, Cotton SC, Harris CA, Gray NM, Little J; TOMBOLA Group (2011). Human papillomavirus infection and anxiety: analyses in women with low-grade cervical cytological abnormalities unaware of their infection status. *PLoS One*. 6(6):e21046. doi:10.1371/journal.pone.0021046 PMID:21698168

Josefsson A, Livak K, Gyllensten U (1999). Detection and quantitation of human papillomavirus by using the fluorescent 5′ exonuclease assay. *J Clin Microbiol*. 37(3):490–6. doi:10.1128/JCM.37.3.490-496.1999 PMID:9986801

Kahn JA, Slap GB, Bernstein DI, Kollar LM, Tissot AM, Hillard PA, et al. (2005). Psychological, behavioral, and interpersonal impact of human papillomavirus and Pap test results. *J Womens Health (Larchmt)*. 14(7):650–9. doi:10.1089/jwh.2005.14.650 PMID:16181021

Karjalainen L, Anttila A, Nieminen P, Luostarinen T, Virtanen A (2016). Self-sampling in cervical cancer screening: comparison of a brush-based and a lavage-based cervicovaginal self-sampling device. *BMC Cancer*. 16(1):221. doi:10.1186/s12885-016-2246-9 PMID:26979237

Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L, et al. (2011). Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncol*. 12(7):663–72. doi:10.1016/S1470-2045(11)70145-0 PMID:21684207

Kaufman HW, Alagia DP, Chen Z, Onisko A, Austin RM (2020). Contributions of liquid-based (Papanicolaou) cytology and human papillomavirus testing in cotesting for detection of cervical cancer and precancer

in the United States. *Am J Clin Pathol*. 154(4):510–6. doi:10.1093/ajcp/aqaa074 PMID:32637991

Kitchener HC, Almonte M, Gilham C, Dowie R, Stoykova B, Sargent A, et al.; ARTISTIC Trial Study Group (2009b). ARTISTIC: a randomised trial of human papillomavirus (HPV) testing in primary cervical screening. *Health Technol Assess*. 13(51):iii–iv, 1–150. doi:10.3310/hta13510 PMID:19891902

Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, et al. (2009a). HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol*. 10(7):672–82. doi:10.1016/S1470-2045(09)70156-1 PMID:19540162

Kitchener HC, Canfell K, Gilham C, Sargent A, Roberts C, Desai M, et al. (2014). The clinical effectiveness and cost-effectiveness of primary human papillomavirus cervical screening in England: extended follow-up of the ARTISTIC randomised trial cohort through three screening rounds. *Health Technol Assess*. 18(23):1–196. doi:10.3310/hta18230 PMID:24762804

Kitchener HC, Fletcher I, Roberts C, Wheeler P, Almonte M, Maguire P (2008). The psychosocial impact of human papillomavirus testing in primary cervical screening – a study within a randomized trial. *Int J Gynecol Cancer*. 18(4):743–8. doi:10.1111/j.1525-1438.2007.01113.x PMID:17944916

Kitchener HC, Gilham C, Sargent A, Bailey A, Albrow R, Roberts C, et al. (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(6):864–71. doi:10.1016/j.ejca.2011.01.008 PMID:21334200

Kocsis A, Takács T, Jeney C, Schaff Z, Koiss R, Járay B, et al. (2017). Performance of a new HPV and biomarker assay in the management of hrHPV positive women: subanalysis of the ongoing multicenter TRACE clinical trial ($n > 6{,}000$) to evaluate POU4F3 methylation as a potential biomarker of cervical precancer and cancer. *Int J Cancer*. 140(5):1119–33. doi:10.1002/ijc.30534 PMID:27874187

Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PPL, Mustafa RA, et al. (2017). Cytology versus HPV testing for cervical cancer screening in the general population. *Cochrane Database Syst Rev*. 8:CD008587. doi:10.1002/14651858.CD008587.pub2 PMID:28796882

Kroenke K, Spitzer RL, Williams JB, Löwe B (2009). An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics*. 50(6):613–21. PMID:19996233

Kulasingam SL, Hughes JP, Kiviat NB, Mao C, Weiss NS, Kuypers JM, et al. (2002). Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. *JAMA*. 288(14):1749–57. doi:10.1001/jama.288.14.1749 PMID:12365959

Kwan TTC, Cheung ANY, Lo SST, Lee PWH, Tam KF, Chan KKL, et al. (2011). Psychological burden of testing positive for high-risk human papillomavirus on women with atypical cervical cytology: a prospective study. *Acta Obstet Gynecol Scand*. 90(5):445–51. doi:10.1111/j.1600-0412.2011.01092.x PMID:21306349

Labani S, Asthana S (2016). Age-specific performance of careHPV versus Papanicolaou and visual inspection of cervix with acetic acid testing in a primary cervical cancer screening. *J Epidemiol Community Health*. 70(1):72–7. doi:10.1136/jech-2015-205851 PMID:26248549

Lamin H, Eklund C, Elfström KM, Carlsten-Thor A, Hortlund M, Elfgren K, et al. (2017). Randomised healthcare policy evaluation of organised primary human papillomavirus screening of women aged 56–60. *BMJ Open*. 7(5):e014788. doi:10.1136/bmjopen-2016-014788 PMID:28566363

Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM; Cochrane Diagnostic Test Accuracy Working Group (2008). Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 149(12):889–97. doi:10.7326/0003-4819-149-12-200812160-00008 PMID:19075208

Leinonen MK, Anttila A, Malila N, Dillner J, Forslund O, Nieminen P (2013). Type- and age-specific distribution of human papillomavirus in women attending cervical cancer screening in Finland. *Br J Cancer*. 109(11):2941–50. doi:10.1038/bjc.2013.647 PMID:24136148

Leinonen MK, Nieminen P, Lönnberg S, Malila N, Hakama M, Pokhrel A, et al. (2012). Detection rates of precancerous and cancerous cervical lesions within one screening round of primary human papillomavirus DNA testing: prospective randomised trial in Finland. *BMJ*. 345:e7789. doi:10.1136/bmj.e7789 PMID:23197596

Leinonen MK, Schee K, Jonassen CM, Lie AK, Nystrand CF, Rangberg A, et al. (2018). Safety and acceptability of human papillomavirus testing of self-collected specimens: a methodologic study of the impact of collection devices and HPV assays on sensitivity for cervical cancer and high-grade lesions. *J Clin Virol*. 99–100:22–30. doi:10.1016/j.jcv.2017.12.008 PMID:29289814

Li N, Shi JF, Franceschi S, Zhang WH, Dai M, Liu B, et al. (2009). Different cervical cancer screening approaches in a Chinese multicentre study. *Br J Cancer*. 100(3):532–7. doi:10.1038/sj.bjc.6604840 PMID:19127262

Longatto-Filho A, Naud P, Derchain SFM, Roteli-Martins C, Tatti S, Hammes LS, et al. (2012). Performance characteristics of Pap test, VIA, VILI, HR-HPV testing, cervicography, and colposcopy in diagnosis of significant cervical pathology. *Virchows Arch*. 460(6):577–85. doi:10.1007/s00428-012-1242-y PMID:22562132

Lorincz AT (1997). Methods of DNA hybridisation and their clinical applicability to human papillomavirus detection. In: Franco EL, Monsonego J, editors. New developments in cervical cancer screening and prevention. Oxford, UK: Blackwell; pp. 325–37.

Luo H, Du H, Belinson JL, Wu R (2019). Evaluation of alternately combining HPV viral load and 16/18 genotyping in secondary screening algorithms. *PLoS One*. 14(7):e0220200. doi:10.1371/journal.pone.0220200 PMID:31348794

Luyten A, Buttmann-Schweiger N, Luyten K, Mauritz C, Reinecke-Lüthge A, Pietralla M, et al. (2014). Early detection of CIN3 and cervical cancer during long-term follow-up using HPV/Pap smear co-testing and risk-adapted follow-up in a locally organised screening programme. *Int J Cancer*. 135(6):1408–16. doi:10.1002/ijc.28783 PMID:24519782

Machalek DA, Roberts JM, Garland SM, Thurloe J, Richards A, Chambers I, et al. (2019). Routine cervical screening by primary HPV testing: early findings in the renewed National Cervical Screening Program. *Med J Aust*. 211(3):113–9. doi:10.5694/mja2.50223 PMID:31168828

Maggino T, Sciarrone R, Murer B, Dei Rossi MR, Fedato C, Maran M, et al. (2016). Screening women for cervical cancer carcinoma with a HPV mRNA test: first results from the Venice pilot program. *Br J Cancer*. 115(5):525–32. doi:10.1038/bjc.2016.216 PMID:27490801

Mahmud SM, Sangwa-Lugoma G, Nasr SH, Kayembe PK, Tozin RR, Drouin P, et al. (2012). Comparison of human papillomavirus testing and cytology for cervical cancer screening in a primary health care setting in the Democratic Republic of the Congo. *Gynecol Oncol*. 124(2):286–91. doi:10.1016/j.ygyno.2011.10.031 PMID:22062546

Maissi E, Marteau TM, Hankins M, Moss S, Legood R, Gray A (2004). Psychological impact of human papillomavirus testing in women with borderline or mildly dyskaryotic cervical smear test results: cross sectional questionnaire study. *BMJ*. 328(7451):1293. doi:10.1136/bmj.328.7451.1293 PMID:15166066

Maissi E, Marteau TM, Hankins M, Moss S, Legood R, Gray A (2005). The psychological impact of human papillomavirus testing in women with borderline or mildly dyskaryotic cervical smear test results: 6-month follow-up. *Br J Cancer*. 92(6):990–4. doi:10.1038/sj.bjc.6602411 PMID:15785734

Malone C, Tiro JA, Buist DSM, Beatty T, Lin J, Kimbel K, et al. (2020). Reactions of women underscreened for cervical cancer who received unsolicited human papillomavirus self-sampling kits. *J Med Screen*. 27(3):146–56. doi:10.1177/0969141319885994 PMID:31744374

Mao C, Kulasingam SL, Whitham HK, Hawes SE, Lin J, Kiviat NB (2017). Clinician and patient acceptability of self-collected human papillomavirus testing for cervical cancer screening. *J Womens Health (Larchmt)*. 26(6):609–15. doi:10.1089/jwh.2016.5965 PMID:28332888

Markovic-Denic L, Djuric O, Maksimovic N, Popovac S, Kesic V (2018). Effects of human papillomavirus awareness and knowledge on psychological state of women referred to cervical cancer screening. *J Low Genit Tract Dis*. 22(3):178–83. doi:10.1097/LGT.0000000000000397 PMID:29672447

Marteau TM, Bekker H (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol*. 31(3):301–6. doi:10.1111/j.2044-8260.1992.tb00997.x PMID:1393159

Mast TC, Zhu X, Demuro-Mercon C, Cummings HW, Sings HL, Ferris DG (2009). Development and psychometric properties of the HPV Impact Profile (HIP) to assess the psychosocial burden of HPV. *Curr Med Res Opin*. 25(11):2609–19. doi:10.1185/03007990903238786 PMID:19739938

Mayrand MH, Duarte-Franco E, Coutlée F, Rodrigues I, Walter SD, Ratnam S, et al.; CCCaST Study Group (2006). Randomized controlled trial of human papillomavirus testing versus Pap cytology in the primary screening for cervical cancer precursors: design, methods and preliminary accrual results of the Canadian Cervical Cancer Screening Trial (CCCaST). *Int J Cancer*. 119(3):615–23. doi:10.1002/ijc.21897 PMID:16572425

Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al.; Canadian Cervical Cancer Screening Trial Study Group (2007). Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med*. 357(16):1579–88. doi:10.1056/NEJMoa071430 PMID:17942871

Maza M, Melendez M, Masch R, Alfaro K, Chacon A, Gonzalez E, et al. (2018). Acceptability of self-sampling and human papillomavirus testing among non-attenders of cervical cancer screening programs in El Salvador. *Prev Med*. 114:149–55. doi:10.1016/j.ypmed.2018.06.017 PMID:29958860

Mbatha JN, Galappaththi-Arachchige HN, Mtshali A, Taylor M, Ndhlovu PD, Kjetland EF, et al. (2017). Self-sampling for human papillomavirus testing among rural young women of KwaZulu-Natal, South Africa. *BMC Res Notes*. 10(1):702. doi:10.1186/s13104-017-3045-3 PMID:29208043

McBride E, Marlow LAV, Forster AS, Ridout D, Kitchener H, Patnick J, et al. (2020). Anxiety and distress following receipt of results from routine HPV primary testing in cervical screening: the Psychological Impact of Primary Screening (PIPS) study. *Int J Cancer*. 146(8):2113–21. doi:10.1002/ijc.32540 PMID:31251820

McCaffery K, Waller J, Forrest S, Cadman L, Szarewski A, Wardle J (2004). Testing positive for human papillomavirus in routine cervical screening: examination

of psychosocial impact. *BJOG*. 111(12):1437–43. doi:10.1111/j.1471-0528.2004.00279.x PMID:15663132

McCaffery K, Waller J, Nazroo J, Wardle J (2006). Social and psychological impact of HPV testing in cervical screening: a qualitative study. *Sex Transm Infect*. 82(2):169–74. doi:10.1136/sti.2005.016436 PMID:16581749

McCaffery KJ, Irwig L, Turner R, Chan SF, Macaskill P, Lewicka M, et al. (2010). Psychosocial outcomes of three triage methods for the management of borderline abnormal cervical smears: an open randomised trial. *BMJ*. 340:b4491. doi:10.1136/bmj.b4491 PMID:20179125

Meijer CJLM, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G, et al. (2009). Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer*. 124(3):516–20. doi:10.1002/ijc.24010 PMID:18973271

Melnikow J, Henderson JT, Burda BU, Senger CA, Durbin S, Weyrich MS (2018). Screening for cervical cancer with high-risk human papillomavirus testing: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA*. 320(7):687–705. doi:10.1001/jama.2018.10400 PMID:30140883

Mesher D, Szarewski A, Cadman L, Cubie H, Kitchener H, Luesley D, et al. (2010). Long-term follow-up of cervical disease in women screened by cytology and HPV testing: results from the HART study. *Br J Cancer*. 102(9):1405–10. doi:10.1038/sj.bjc.6605619 PMID:20354519

Mittal S, Mandal R, Banerjee D, Das P, Ghosh I, Panda C, et al. (2016). HPV detection-based cervical cancer screening program in low-resource setting: lessons learnt from a community-based demonstration project in India. *Cancer Causes Control*. 27(3):351–8. doi:10.1007/s10552-015-0708-z PMID:26712612

Moberg M, Gustavsson I, Gyllensten U (2004). Type-specific associations of human papillomavirus load with risk of developing cervical carcinoma *in situ*. *Int J Cancer*. 112(5):854–9. doi:10.1002/ijc.20480 PMID:15386375

Modibbo F, Iregbu KC, Okuma J, Leeman A, Kasius A, de Koning M, et al. (2017). Randomized trial evaluating self-sampling for HPV DNA based tests for cervical cancer screening in Nigeria. *Infect Agent Cancer*. 12(1):11. doi:10.1186/s13027-017-0123-z PMID:28184239

Monsonego J, Hudgens MG, Zerat L, Zerat JC, Syrjänen K, Halfon P, et al. (2011). Evaluation of oncogenic human papillomavirus RNA and DNA tests with liquid-based cytology in primary cervical cancer screening: the FASE study. *Int J Cancer*. 129(3):691–701. doi:10.1002/ijc.25726 PMID:20941740

Moy LM, Zhao FH, Li LY, Ma JF, Zhang QM, Chen F, et al. (2010). Human papillomavirus testing and cervical cytology in primary screening for cervical cancer among women in rural China: comparison of sensitivity, specificity, and frequency of referral. *Int J Cancer*. 127(3):646–56. doi:10.1002/ijc.25071 PMID:19960441

Muangto T, Chanthasenanont A, Lertvutivivat S, Nanthakomon T, Pongrojpaw D, Bhamarapravatana K, et al. (2016). Experience of combined liquid based cervical cytology and high-risk HPV mRNA for cervical cancer screening in Thammasat University Hospital. *Asian Pac J Cancer Prev*. 17(9):4409–13. PMID:27797253

Mustafa RA, Santesso N, Khatib R, Mustafa AA, Wiercioch W, Kehar R, et al. (2016). Systematic reviews and meta-analyses of the accuracy of HPV tests, visual inspection with acetic acid, cytology, and colposcopy. *Int J Gynaecol Obstet*. 132(3):259–65. doi:10.1016/j.ijgo.2015.07.024 PMID:26851054

Muwonge R, Wesley RS, Nene BM, Shastri SS, Jayant K, Malvi SG, et al. (2014). Evaluation of cytology and visual triage of human papillomavirus-positive women in cervical cancer prevention in India. *Int J Cancer*. 134(12):2902–9. doi:10.1002/ijc.28627 PMID:24272364

Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K, et al. (2007). Human papillomavirus and Papanicolaou tests to screen for cervical cancer. [Erratum in: N Engl J Med. 2008 359(15):1637. Johansson, Bo (added)] *N Engl J Med*. 357(16):1589–97. doi:10.1056/NEJMoa073204 PMID:17942872

Naucler P, Ryd W, Törnberg S, Strand A, Wadell G, Elfgren K, et al. (2009). Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst*. 101(2):88–99. doi:10.1093/jnci/djn444 PMID:19141778

Nelson EJ, Maynard BR, Loux T, Fatla J, Gordon R, Arnold LD (2017). The acceptability of self-sampled screening for HPV DNA: a systematic review and meta-analysis. *Sex Transm Infect*. 93(1):56–61. doi:10.1136/sextrans-2016-052609 PMID:28100761

Nieves L, Enerson CL, Belinson S, Brainard J, Chiesa-Vottero A, Nagore N, et al. (2013). Primary cervical cancer screening and triage using an mRNA human papillomavirus assay and visual inspection. *Int J Gynecol Cancer*. 23(3):513–8. doi:10.1097/IGC.0b013e318280f3bc PMID:23334437

O'Connor M, Costello L, Murphy J, Prendiville W, Martin CM, O'Leary JJ, et al.; Irish Screening Research Consortium (CERVIVA) (2014). 'I don't care whether it's HPV or ABC, I just want to know if I have cancer.' Factors influencing women's emotional responses to undergoing human papillomavirus testing in routine management in cervical screening: a qualitative study. *BJOG*. 121(11):1421–9. doi:10.1111/1471-0528.12741 PMID:24690225

Ogilvie GS, Krajden M, van Niekerk D, Smith LW, Cook D, Ceballos K, et al. (2017). HPV For Cervical Cancer Screening (HPV FOCAL): complete round 1 results of a randomized trial comparing HPV-based primary screening to liquid-based cytology for cervical cancer. *Int J Cancer*. 140(2):440–8. doi:10.1002/ijc.30454 PMID:27685757

Ogilvie GS, van Niekerk D, Krajden M, Smith LW, Cook D, Gondara L, et al. (2018). Effect of screening with primary cervical HPV testing vs cytology testing on high-grade cervical intraepithelial neoplasia at 48 months: the HPV FOCAL randomized clinical trial. *JAMA*. 320(1):43–52. doi:10.1001/jama.2018.7464 PMID:29971397

Oketch SY, Kwena Z, Choi Y, Adewumi K, Moghadassi M, Bukusi EA, et al. (2019). Perspectives of women participating in a cervical cancer screening campaign with community-based HPV self-sampling in rural western Kenya: a qualitative study. *BMC Women's Health*. 19(1):75. doi:10.1186/s12905-019-0778-2 PMID: 31196175

Ovestad IT, Vennestrøm U, Andersen L, Gudlaugsson E, Munk AC, Malpica A, et al. (2011). Comparison of different commercial methods for HPV detection in follow-up cytology after ASCUS/LSIL, prediction of CIN2–3 in follow up biopsies and spontaneous regression of CIN2–3. *Gynecol Oncol*. 123(2):278–83. doi:10.1016/j.ygyno.2011.07.024 PMID:21835440

Pan Q, Belinson JL, Li L, Pretorius RG, Qiao YL, Zhang WH, et al. (2003). A thin-layer, liquid-based Pap test for mass screening in an area of China with a high incidence of cervical carcinoma. A cross-sectional, comparative study. *Acta Cytol*. 47(1):45–50. doi:10.1159/000326474 PMID:12585030

Pan QJ, Hu SY, Guo HQ, Zhang WH, Zhang X, Chen W, et al. (2014). Liquid-based cytology and human papillomavirus testing: a pooled analysis using the data from 13 population-based cervical cancer screening studies from China. *Gynecol Oncol*. 133(2):172–9. doi:10.1016/j.ygyno.2014.03.008 PMID:24631450

Papa D, Moore Simas TA, Reynolds M, Melnitsky H (2009). Assessing the role of education in women's knowledge and acceptance of adjunct high-risk human papillomavirus testing for cervical cancer screening. *J Low Genit Tract Dis*. 13(2):66–71. doi:10.1097/LGT.0b013e31818a53f0 PMID:19387125

Pasquale L, Giorgi Rossi P, Carozzi F, Pedretti C, Ruggeri C, Scalvinoni V, et al. (2015). Cervical cancer screening with HPV testing in the Valcamonica (Italy) screening programme. *J Med Screen*. 22(1):38–48. doi:10.1177/0969141314561707 PMID:25431452

Passamonti B, Gustinucci D, Giorgi Rossi P, Cesarini E, Bulletti S, Carlani A, et al. (2017). Cervical human papilloma virus (HPV) DNA primary screening test: results of a population-based screening programme in central Italy. *J Med Screen*. 24(3):153–62. doi:10.1177/0969141316663580 PMID:27614992

Patel H, Moss EL, Sherman SM (2018). HPV primary cervical screening in England: women's awareness and attitudes. *Psychooncology*. 27(6):1559–64. doi:10.1002/pon.4694 PMID:29521462

Petry KU, Menton S, Menton M, van Loenen-Frosch F, de Carvalho Gomes H, Holz B, et al. (2003). Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *Br J Cancer*. 88(10):1570–7. doi:10.1038/sj.bjc.6600918 PMID:12771924

Poljak M, Oštrbenk Valenčak A, Gimpelj Domjanič G, Xu L, Arbyn M (2020). Commercially available molecular tests for human papillomaviruses: a global overview. *Clin Microbiol Infect*. 26(9):1144–50. doi:10.1016/j.cmi.2020.03.033 PMID:32247892

Polman NJ, de Haan Y, Veldhuijzen NJ, Heideman DAM, de Vet HCW, Meijer CJLM, et al. (2019c). Experience with HPV self-sampling and clinician-based sampling in women attending routine cervical screening in the Netherlands. *Prev Med*. 125:5–11. doi:10.1016/j.ypmed.2019.04.025 PMID:31054907

Polman NJ, Ebisch RMF, Heideman DAM, Melchers WJG, Bekkers RLM, Molijn AC, et al. (2019b). Performance of human papillomavirus testing on self-collected versus clinician-collected samples for the detection of cervical intraepithelial neoplasia of grade 2 or worse: a randomised, paired screen-positive, non-inferiority trial. *Lancet Oncol*. 20(2):229–38. doi:10.1016/S1470-2045(18)30763-0 PMID:30658933

Polman NJ, Veldhuijzen NJ, Heideman DAM, Snijders PJF, Meijer CJLM, Berkhof J (2017). HPV-positive women with normal cytology remain at increased risk of CIN3 after a negative repeat HPV test. *Br J Cancer*. 117(10):1557–61. doi:10.1038/bjc.2017.309 PMID:28881359

Polman NJ, Veldhuijzen NJ, Heideman DAM, Snijders PJF, Meijer CJLM, Berkhof J (2019a). Management of HPV-positive women in cervical screening using results from two consecutive screening rounds. *Int J Cancer*. 144(9):2339–46. doi:10.1002/ijc.32004 PMID: 30565673

Qiao YL, Jeronimo J, Zhao FH, Schweizer J, Chen W, Valdez M, et al. (2014). Lower cost strategies for triage of human papillomavirus DNA-positive women. *Int J Cancer*. 134(12):2891–901. doi:10.1002/ijc.28616 PMID:24248915

Rebolj M, Bonde J, Ejegod D, Preisler S, Rygaard C, Lynge E (2015). A daunting challenge: human papillomavirus assays and cytology in primary cervical screening of women below age 30 years. *Eur J Cancer*. 51(11):1456–66. doi:10.1016/j.ejca.2015.04.012 PMID:25979832

Rebolj M, Rimmer J, Denton K, Tidy J, Mathews C, Ellis K, et al. (2019). Primary cervical screening with high risk human papillomavirus testing: observational study. *BMJ*. 364:l240. doi:10.1136/bmj.l240 PMID:30728133

Rezhake R, Hu SY, Zhao S, Xu XQ, Zhao XL, Zhang L, et al. (2018). Eight-type human papillomavirus E6/E7 oncoprotein detection as a novel and promising triage strategy for managing HPV-positive women. *Int J Cancer*. 144(1):34–42. doi:10.1002/ijc.31633 PMID:29943809

Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkmans NWJ, Heideman DAM, et al. (2012a). Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol*. 13(1):78–88. doi:10.1016/S1470-2045(11)70296-0 PMID:22177579

Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VMH, Rozendaal L, Heideman DAM, et al. (2012b). HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. *Br J Cancer*. 106(5):975–81. doi:10.1038/bjc.2011.581 PMID:22251922

Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Hesselink AT, Rozendaal L, et al. (2012c). Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. *Int J Cancer*. 130(3):602–10. doi:10.1002/ijc.26056 PMID:21400507

Ronco G, Brezzi S, Carozzi F, Dalla Palma P, Giorgi-Rossi P, Minucci D, et al.; NTCC study group (2007a). The New Technologies for Cervical Cancer Screening randomised controlled trial. An overview of results during the first phase of recruitment. *Gynecol Oncol*. 107(Suppl 1):S230–2. doi:10.1016/j.ygyno.2007.07.021 PMID:17822751

Ronco G, Cuzick J, Segnan N, Brezzi S, Carozzi F, Folicaldi S, et al.; NTCC working group (2007b). 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*. 43(3):476–80. doi:10.1016/j.ejca.2006.11.013 PMID:17223540

Ronco G, Dillner J, Elfström KM, Tunesi S, Snijders PJF, Arbyn M, et al.; International HPV screening working group (2014). Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet*. 383(9916):524–32. doi:10.1016/S0140-6736(13)62218-7 PMID:24192252

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al.; New Technologies for Cervical Cancer screening (NTCC) Working Group (2010). Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol*. 11(3):249–57. doi:10.1016/S1470-2045(09)70360-2 PMID:20089449

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al.; New Technologies for Cervical Cancer Screening Working Group (2008). 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*. 100(7):492–501. doi:10.1093/jnci/djn065 PMID:18364502

Ronco G, Giorgi-Rossi P, Carozzi F, Dalla Palma P, Del Mistro A, De Marco L, et al.; New Technologies for Cervical Cancer Screening Working Group (2006a). Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. *Lancet Oncol*. 7(7):547–55. doi:10.1016/S1470-2045(06)70731-8 PMID:16814206

Ronco G, Segnan N, Giorgi-Rossi P, Zappa M, Casadei GP, Carozzi F, et al.; New Technologies for Cervical Cancer Working Group (2006b). Human papillomavirus testing and liquid-based cytology: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. *J Natl Cancer Inst*. 98(11):765–74. doi:10.1093/jnci/djj209 PMID:16757701

Salmerón J, Lazcano-Ponce E, Lorincz A, Hernández M, Hernández P, Leyva A, et al. (2003). Comparison of HPV-based assays with Papanicolaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes Control*. 14(6):505–12. doi:10.1023/A:1024806707399 PMID:12948281

Sangrajrang S, Laowahutanont P, Wongsena M, Muwonge R, Karalak A, Imsamran W, et al. (2017). Comparative accuracy of Pap smear and HPV screening in Ubon Ratchathani in Thailand. *Papillomavirus Res*. 3:30–5. doi:10.1016/j.pvr.2016.12.004 PMID:28720454

Sankaranarayanan R, Chatterji R, Shastri SS, Wesley RS, Basu P, Mahe C, et al. (2004). Accuracy of human papillomavirus testing in primary screening of cervical neoplasia: results from a multicenter study in India. *Int J Cancer*. 112(2):341–7. doi:10.1002/ijc.20396 PMID:15352050

Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS, et al.; Osmanabad District Cervical Screening Study Group (2005). A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. *Int J Cancer*. 116(4):617–23. doi:10.1002/ijc.21050 PMID:15818610

Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. (2009). HPV screening for cervical cancer in rural India. *N Engl J Med*. 360(14):1385–94. doi:10.1056/NEJMoa0808516 PMID:19339719

Sargent A, Fletcher S, Bray K, Kitchener HC, Crosbie EJ (2019). Cross-sectional study of HPV testing in self-sampled urine and comparison with matched vaginal and cervical samples in women attending colposcopy for

the management of abnormal cervical screening. *BMJ Open*. 9(4):e025388. doi:10.1136/bmjopen-2018-025388 PMID:31036707

Sarian LO, Derchain S, Shabalova I, Tatti S, Naud P, Longatto-Filho A, et al.; members of the Latin American Screening Study (LAMS) and the New Independent States of the Former Soviet Union (NIS) groups (2010). Optional screening strategies for cervical cancer using standalone tests and their combinations among low- and medium-income populations in Latin America and Eastern Europe. *J Med Screen*. 17(4):195–203. doi:10.1258/jms.2010.010016 PMID:21258130

Schiffman M, Kinney WK, Cheung LC, Gage JC, Fetterman B, Poitras NE, et al. (2018). Relative performance of HPV and cytology components of cotesting in cervical screening. *J Natl Cancer Inst*. 110(5):501–8. doi:10.1093/jnci/djx225 PMID:29145648

Senkomago V, Ting J, Kwatampora J, Gukare H, Mugo N, Kimani J, et al. (2018). High-risk HPV-RNA screening of physician- and self-collected specimens for detection of cervical lesions among female sex workers in Nairobi, Kenya. *Int J Gynaecol Obstet*. 143(2):217–24. doi:10.1002/ijgo.12628 PMID:30047987

Sherman ME, Schiffman M, Cox JT; Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study Group (2002). 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*. 94(2):102–7. doi:10.1093/jnci/94.2.102 PMID:11792748

Snijders PJF, van den Brule AJC, Meijer CJLM (2003). The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. *J Pathol*. 201(1):1–6. doi:10.1002/path.1433 PMID:12950011

Stoler MH, Wright TC Jr, Cuzick J, Dockter J, Reid JL, Getman D, et al. (2013). APTIMA HPV assay performance in women with atypical squamous cells of undetermined significance cytology results. *Am J Obstet Gynecol*. 208(2):144.e1–8. doi:10.1016/j.ajog.2012.12.003 PMID:23220509

Sultana F, Mullins R, English DR, Simpson JA, Drennan KT, Heley S, et al. (2015). Women's experience with home-based self-sampling for human papillomavirus testing. *BMC Cancer*. 15(1):849. doi:10.1186/s12885-015-1804-x PMID:26536865

Tang NS, Tang ML, Chan ISF (2003). On tests of equivalence via non-unity relative risk for matched-pair design. *Stat Med*. 22(8):1217–33. doi:10.1002/sim.1213 PMID:12687652

Terrazas S, Ibáñez C, Lagos M, Poggi H, Brañes J, Barriga MI, et al. (2015). Human papillomavirus testing in cervical cancer screening at a public health service of Santiago, Chile. [in Spanish] *Rev Med Chil*. 143(1):56–62. doi:10.4067/S0034-98872015000100007 PMID:25860269

Thomsen LT, Kjær SK, Munk C, Frederiksen K, Ørnskov D, Waldstrøm M (2020). Clinical performance of human papillomavirus (HPV) testing versus cytology for cervical cancer screening: results of a large Danish implementation study. *Clin Epidemiol*. 12:203–13. doi:10.2147/CLEP.S243546 PMID:32110112

Tian Q, Li Y, Wang F, Li Y, Xu J, Shen Y, et al. (2014). MicroRNA detection in cervical exfoliated cells as a triage for human papillomavirus-positive women. *J Natl Cancer Inst*. 106(9):dju241. doi:10.1093/jnci/dju241 PMID:25190727

Torres-Ibarra L, Cuzick J, Lorincz AT, Spiegelman D, Lazcano-Ponce E, Franco EL, et al.; FRIDA Study Group (2019). Comparison of HPV-16 and HPV-18 genotyping and cytological testing as triage testing within human papillomavirus-based screening in Mexico. *JAMA Netw Open*. 2(11):e1915781. doi:10.1001/jamanetworkopen.2019.15781 PMID:31747033

Tshomo U, Franceschi S, Tshokey T, Tobgay T, Baussano I, Tenet V, et al. (2017). Evaluation of cytology versus human papillomavirus-based cervical cancer screening algorithms in Bhutan. *Oncotarget*. 8(42):72438–46. doi:10.18632/oncotarget.19783 PMID:29069800

Veijalainen O, Kares S, Kujala P, Vuento R, Osuala V, Tirkkonen M, et al. (2019). Implementation of HPV-based cervical cancer screening in an organised regional screening programme: 3 years of experience. *Cytopathology*. 30(2):150–6. doi:10.1111/cyt.12652 PMID:30421573

Veldhuijzen NJ, Polman NJ, Snijders PJF, Meijer CJLM, Berkhof J (2017). Stratifying HPV-positive women for CIN3+ risk after one and two rounds of HPV-based screening. *Int J Cancer*. 141(8):1551–60. doi:10.1002/ijc.30865 PMID:28670823

Verdoodt F, Szarewski A, Halfon P, Cuschieri K, Arbyn M (2013). Triage of women with minor abnormal cervical cytology: meta-analysis of the accuracy of an assay targeting messenger ribonucleic acid of 5 high-risk human papillomavirus types. *Cancer Cytopathol*. 121(12):675–87. doi:10.1002/cncy.21325 PMID:23881840

von Karsa L, Arbyn M, De Vuyst H, Dillner J, Dillner L, Franceschi S, et al. (2015). European guidelines for quality assurance in cervical cancer screening. Summary of the supplements on HPV screening and vaccination. *Papillomavirus Res*. 1:22–31. doi:10.1016/j.pvr.2015.06.006

Waller J, Marlow LAV, Wardle J (2007). The association between knowledge of HPV and feelings of stigma, shame and anxiety. *Sex Transm Infect*. 83(2):155–9. doi:10.1136/sti.2006.023333 PMID:17098767

Wang MZ, Feng RM, Wang S, Duan XZ, Li D, Zhang X, et al. (2019). Clinical performance of human papillomavirus testing and visual inspection with acetic acid in

primary, combination, and sequential cervical cancer screening in China. *Sex Transm Dis*. 46(8):540–7. doi:10.1097/OLQ.0000000000001026 PMID:31295223

Wang M, Hu S, Zhao S, Zhang W, Pan Q, Zhang X, et al. (2017). Accuracy of triage strategies for human papillomavirus DNA-positive women in low-resource settings: a cross-sectional study in China. *Chin J Cancer Res*. 29(6):496–509. doi:10.21147/j.issn.1000-9604.2017.06.04 PMID:29353972

Wang SM, Shi JF, Kang DJ, Song P, Qiao YL; Chinese HPV Study Group (2011). Impact of human papillomavirus-related lesions on quality of life: a multicenter hospital-based study of women in mainland China. *Int J Gynecol Cancer*. 21(1):182–8. doi:10.1097/IGC.0b013e3181ffbed8 PMID:21330842

Wardle J, Pernet A, Stephens D (1995). Psychological consequences of positive results in cervical cancer screening. *Psychol Health*. 10(3):185–94. doi:10.1080/08870449508401948

Wentzensen N, Fetterman B, Castle PE, Schiffman M, Wood SN, Stiemerling E, et al. (2015). p16/Ki-67 dual stain cytology for detection of cervical precancer in HPV-positive women. *J Natl Cancer Inst*. 107(12):djv257. doi:10.1093/jnci/djv257 PMID:26376685

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al.; QUADAS-2 Group (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 155(8):529–36. doi:10.7326/0003-4819-155-8-201110180-00009 PMID:22007046

WHO (2010). Human papillomavirus laboratory manual. 1st ed. Geneva, Switzerland: World Health Organization, Department of Immunization, Vaccines and Biologicals. Available from: https://apps.who.int/iris/handle/10665/70505.

WHO (2019). Protocol for the performance evaluation of qualitative HPV nucleic acid based assays for WHO prequalification assessment. Geneva, Switzerland: World Health Organization.

WHO (2020a). WHO technical guidance and specifications of medical devices for screening and treatment of precancerous lesions in the prevention of cervical cancer. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/331698.

WHO (2020b). Global strategy to accelerate the elimination of cervical cancer as a public health problem. Geneva, Switzerland: World Health Organization. Available from: https://apps.who.int/iris/handle/10665/336583.

Wright TC, Stoler MH, Behrens CM, Sharma A, Zhang G, Wright TL (2015). Primary cervical cancer screening with human papillomavirus: end of study results from the ATHENA study using HPV as the first-line screening test. *Gynecol Oncol*. 136(2):189–97. doi:10.1016/j.ygyno.2014.11.076 PMID:25579108

Wu R, Belinson SE, Du H, Na W, Qu X, Wu R, et al. (2010). Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. *Int J Gynecol Cancer*. 20(8):1411–4. PMID:21051986

Wu Q, Zhao X, Fu Y, Wang X, Zhang X, Tian X, et al. (2017). A cross-sectional study on HPV testing with type 16/18 genotyping for cervical cancer screening in 11,064 Chinese women. *Cancer Med*. 6(5):1091–101. doi:10.1002/cam4.1060 PMID:28378404

Zhang L, Xu XQ, Hu SY, Chen F, Zhang X, Pan QJ, et al. (2018). Durability of clinical performance afforded by self-collected HPV testing: a 15-year cohort study in China. *Gynecol Oncol*. 151(2):221–8. doi:10.1016/j.ygyno.2018.09.012 PMID:30269870

Zhao FH, Jeronimo J, Qiao YL, Schweizer J, Chen W, Valdez M, et al. (2013). An evaluation of novel, lower-cost molecular screening tests for human papillomavirus in rural China. *Cancer Prev Res (Phila)*. 6(9):938–48. doi:10.1158/1940-6207.CAPR-13-0091 PMID:23878179

Zhao FH, Lin MJ, Chen F, Hu SY, Zhang R, Belinson JL, et al.; Cervical Cancer Screening Group in China (2010). 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*. 11(12):1160–71. doi:10.1016/S1470-2045(10)70256-4 PMID:21075054

Zhao YQ, Dai Y, Dang L, Kong LH, Zhang Y, Feng RM, et al. (2018). Real-world research on cervical cancer screening program and effect evaluation for Chinese population. [in Chinese] *Chin J Oncol*. 40(10):764–71. doi:10.3760/cma.j.issn.0253-3766.2018.10.008 PMID:30392341

Zhao J, Du H, Belinson JL, Qu X, Zhang W, Mei J, et al. (2016b). Evaluation of the Cervista HPV A9 group in screening patients for cervical cancer. *J Med Screen*. 23(1):38–43. doi:10.1177/0969141315604863 PMID:26466824

Zhao FH, Hu SY, Zhang Q, Zhang X, Pan QJ, Zhang WH, et al. (2016a). Risk assessment to guide cervical screening strategies in a large Chinese population. *Int J Cancer*. 138(11):2639–47. doi:10.1002/ijc.30012 PMID:26800481

Zigmond AS, Snaith RP (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 67(6):361–70. doi:10.1111/j.1600-0447.1983.tb09716.x PMID:6880820

Zorzi M, Del Mistro A, Giorgi Rossi P, Laurino L, Battagello J, Lorio M, et al. (2020). Risk of CIN2 or more severe lesions after negative HPV-mRNA E6/E7 overexpression assay and after negative HPV-DNA test: concurrent cohorts with a 5-year follow-up. *Int J Cancer*. 146(11):3114–23. doi:10.1002/ijc.32695 PMID:31583683

Zorzi M, Frayle H, Rizzi M, Fedato C, Rugge M, Penon MG, et al.; Veneto HPV-screening Working Group (2017). A 3-year interval is too short for re-screening women testing negative for human papillomavirus: a population-based cohort study. *BJOG*. 124(10):1585–93. doi:10.1111/1471-0528.14575 PMID:28120382

zur Hausen H (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer*. 2(5):342–50. doi:10.1038/nrc798 PMID:12044010

4.5 Colposcopy

A colposcope is a low-magnification, light-illuminated, stereoscopic, binocular field microscope. It is used for visual examination of the lower genital tract, including the cervix. Colposcopic examination facilitates the identification of the TZ (see Fig. 1.18 in Section 1.2.5), which is where most cervical cancers originate, and the characterization and localization of intraepithelial lesions in the lower genital tract to guide biopsies, where necessary, for confirmation of disease status.

In the 20th century, colposcopy was used in many countries as part of a standard gynaecological examination (van Niekerk et al., 1998). It is still used as a primary screening tool, together with cytology, by some clinicians in a few countries in Europe and Latin America. The rationale for this combined testing approach is that the use of the colposcope to guide cytology sample collection may decrease the false-negative and false-positive rates associated with blind sampling, and may also reduce the need for women to be recalled for repeat cytology (van Niekerk et al., 1998). However, there is no agreement about whether colposcopic impression improves the quality of cytology testing (Hilgarth & Menton, 1996; Schulmeyer et al., 2020). Moreover, it has been shown that colposcopy does not perform well for primary screening (Leeson et al., 2014; AEPCC, 2018). In contrast, there is wide consensus that colposcopy is the cornerstone of management of women with a positive Pap test result or symptomatic women. Table 4.33 shows the indications for performing colposcopy.

4.5.1 Technical description of a colposcopic examination

In 1925, Hinselmann (Hinselmann, 1925; Jordan, 1985) designed the colposcope and described how to enhance the colposcopic view of the cervical epithelium to recognize cervical cancer and precancer by staining the cervix with acetic acid (Soutter, 1993). In 1929, Schiller introduced the use of iodine and showed that areas of the cervix harbouring early cervical cancer did not stain with iodine, in contrast to the dark staining of normal squamous epithelium of the ectocervix (Schiller, 1933; Colgan & Lickrish, 1990; Bappa & Yakasai, 2013). Initially colposcopy was used for primary screening, but during the 1960s studies showed that colposcopy enabled the more accurate localization of suspected lesions after cytology testing, which made it possible to more accurately select biopsy sites and reduced the need for diagnostic conization (Beller & Khatamee, 1966; Ruiz Moreno, 2010). These studies established the basis for the current use of colposcopy within the cytology–colposcopy–histology sequence.

When colposcopy is performed in a competent and quality-assured service, it is a comprehensive examination and provides information that is crucial for optimal clinical management. Colposcopy has important advantages, particularly for women with endocervical or glandular disease, very large lesions, or suspicion of invasion or microinvasive disease, and for lesions that are present during pregnancy or for residual or recurrent disease after treatment.

A colposcopic examination aims to:

- determine the adequacy of the examination;
- determine the site, size, and type of the TZ;
- recognize intraepithelial abnormality where present;
- identify the most accurate biopsy site for sampling; and
- facilitate precise treatment.

Table 4.33 Indications for performing colposcopy

Abnormal results in screening tests (cytology or HPV test) suggesting an increased risk of cervical intraepithelial neoplasia
Follow-up of patients with an intraepithelial lesion before or after treatment
Excisional treatment of premalignant lesions of the cervix, as an auxiliary method to guide the procedure
Presence of clinically apparent leukoplakia or any suspicious-looking or abnormal-looking cervix in the gynaecological examination
Presence of symptoms suggesting cervical cancer (unusual bleeding, abnormal vaginal discharge, etc.)

HPV, human papillomavirus.
Compiled by the Working Group.

(a) The colposcope

A colposcope has the following features (for more details, see Prendiville & Sankaranarayanan, 2017):

- A support for the colposcope head, which is the working part. This support can be either a simple vertical stand that is positioned between the operator's legs or an adjustable horizontal arm connected to a weighted stand that is positioned lateral to the patient and the operator and is attached to the colposcope head by a universal joint.
- Binocular view, so that depth of field may be appreciated. (Improving image-capture systems may reduce the disadvantages of monocular devices.) Depth of field is crucial for accurate assessment of the TZ or when performing excision of the TZ.
- Variable magnification, either stepwise or using a zoom facility.
- White light from a halogen light or, preferably, a light-emitting diode (LED) lamp.
- A green or blue filter, or green or blue light.
- Image capture.
- Facility to adjust the eyepieces to the operator's interpupillary distance.
- Fine focus adjustment.

(b) Performing a colposcopic examination

For a colposcopic examination to be performed competently, the following are required: a well-trained colposcopist, a well-equipped examination room (see Prendiville & Sankaranarayanan, 2017), and a skilled attendant.

The examiner inserts a speculum to expose the cervix and position it in a plane perpendicular to the colposcopic line of vision. The colposcope enables the examination of the whole lower genital tract, including the cervix, vagina, and vulva. The examiner first assesses whether the examination can be performed adequately (Bornstein et al., 2012). If so, the next step is to examine the cervix at low-power magnification and gently cleanse it with saline. The hormonal status and degree of inflammation are assessed. Once adequacy has been confirmed, the TZ is examined at low-power magnification, perhaps with a green filter, before 3% to 5% acetic acid is applied. Use of an endocervical forceps (preferably the Desjardins or Kurihara forceps) is often needed to achieve full visualization of the upper limit of the TZ, particularly in postmenopausal women. Examination of the TZ is performed at both low-power and high-power magnification. Documentation of the examination findings completes the colposcopy, and a management plan may be discussed with the patient.

Table 4.34 Modifications in colposcopic terminology over time

Terminology (name, year)	Normal findings	Abnormal findings	Other terms	Reference
Hinselmann, 1933	Thick leukoplakia	Mosaic leukoplakia	Cervico-uterine ectopy	ASCCP guidelines, Mayeaux & Cox (2013)
Coppleson, 1960	Grade I: not suspicious, white semi-transparent epithelium, flat, with indistinct borders	Grade II: suspicious white epithelium Grade III: opaque epithelium with very suspicious defined borders	Transformation zone	Reid & Campion (1989)
IFCPC Graz, 1975	Normal colposcopy	Atypical transformation zone	Colposcopy not satisfactory Miscellaneous	Stafl (1976)
Reid score, 1985	Category 1: benign, minor dysplasia	Category 2: intermediate Category 3: suspicious	Four criteria: border, colour, vessels, iodine uptake	Reid & Campion (1989)
IFCPC Rome, 1990	Normal colposcopy Cylindrical epithelium: ectopy	Abnormal colposcopy within or outside the transformation zone Fine or coarse mosaic or punctation	Miscellaneous not acetowhite	Stafl & Wilbanks (1991)
IFCPC Barcelona, 2002	Type 1, 2, 3 transformation zone	Minor or major changes Suggestive of low-grade or high-grade lesion	Colposcopy suggestive of invasive cancer	Walker et al. (2003)
IFCPC Rio de Janeiro, 2011	Includes metaplasia and deciduosis	Grade 1 or grade 2 changes Location of lesion, number of cervical quadrants the lesion covers New signs: inner border sign and ridge sign	Includes description of vaginal lesions Incorporates types of excision	Bornstein et al. (2012)

ASCCP, American Society for Colposcopy and Cervical Pathology; IFCPC, International Federation of Cervical Pathology and Colposcopy.

(c) Colposcopic terminology and correlation with histological diagnosis

Different classifications have been used throughout the 90-year history of colposcopy (AEPCC, 2018). Table 4.34 shows the most relevant and clinically used global colposcopic classifications and the modifications that have been introduced over time. Currently, the classification that is most commonly used in health-care practice worldwide is that adopted unanimously by the International Federation of Cervical Pathology and Colposcopy (IFCPC). The most recent IFCPC terminology, prepared in 2011 (Bornstein et al., 2012), is summarized in Table 4.35. However, in this section, results from scientific publications are presented according to the terminology as reported originally, wherever possible.

Substantial information is available on the correlation between the categorization of lesions using the IFCPC classification and the histological diagnosis. Some studies have reported a good correlation between the colposcopic impression and the final diagnosis (Ferris & Litaker, 2005). Some particular findings (such as coarse punctation, coarse mosaic or dense acetowhitening, inner border sign, and ridge sign) have been shown to have a good predictive accuracy for HSIL+/CIN2+ (Vercellino et al., 2013; Beyer et al., 2017; Li et al., 2017), although the sensitivity of colposcopic impression for detection of HSIL+/CIN2+ ranged from 20% to 100%

Table 4.35 2011 IFCPC colposcopic terminology of the cervix

Section	Pattern	
General assessment	Adequate or inadequate; if inadequate, for what reason (e.g. cervix obscured by inflammation, bleeding, scar) Squamocolumnar junction visibility: completely visible, partially visible, not visible Transformation zone types 1, 2, 3	
Normal colposcopic findings	Original squamous epithelium: mature, atrophic Columnar epithelium; ectopy or ectropion Metaplastic squamous epithelium; nabothian cysts; crypt (gland) openings Deciduosis in pregnancy	
Abnormal colposcopic findings	*General principles* Location of the lesion: • Inside or outside the transformation zone • By the "clock position" *Grade 1 (minor)* • Fine mosaic; fine punctation • Thin acetowhite epithelium • Irregular, geographical border	Size of the lesion: • Number of cervical quadrants the lesion covers • Size of the lesion as a percentage of the cervix *Grade 2 (major)* • Sharp border; inner border sign; ridge sign • Dense acetowhite epithelium • Coarse mosaic; coarse punctation • Rapid appearance of acetowhitening • Cuffed crypt (gland) openings
	Non-specific • Leukoplakia (keratosis, hyperkeratosis); erosion • Lugol's staining (Schiller test): stained or non-stained	
Suspicious for invasion	Atypical vessels Additional signs: • Fragile vessels • Irregular surface • Exophytic lesion • Necrosis • Ulceration (necrotic) • Tumour or gross neoplasm	
Miscellaneous findings	Congenital transformation zone Condyloma Polyp (ectocervical or endocervical) Inflammation	Stenosis Congenital anomaly Post-treatment consequence Endometriosis
Excision treatment types	Excision types 1, 2, 3	
Excision specimen dimensions	Length: the distance from the distal or external margin to the proximal or internal margin Thickness: the distance from the stromal margin to the surface of the excised specimen Circumference (optional): the perimeter of the excised specimen	

IFCPC, International Federation of Cervical Pathology and Colposcopy.
From Bornstein et al. (2012).

and the specificity from 96% to 99%. However, some authors have suggested that the degree of concordance depends mainly on the training and the experience or expertise of the colposcopist (Mayeaux & Cox, 2013; American Society for Colposcopy and Cervical Pathology [ASCCP] guidelines, Perkins et al, 2020). High-quality training and quality assurance programmes are essential for the competent practice of colposcopy.

Some attempts have been made to quantify qualitative descriptions into scoring systems, such as the Reid Colposcopic Index (RCI) (Reid & Scalzi, 1985) and the Swede score (Strander et al., 2005). It has been suggested that colposcopic

findings are best assessed formally using a scoring system (Prendiville & Sankaranarayanan, 2017; Ranga et al., 2017; Alan et al., 2020; Schulmeyer et al., 2020). However, some studies report better correlation of histology with colposcopic impression than with colposcopy-based quantitative scores. Li et al. (2017) compared the performance of the IFCPC colposcopic terminology, the RCI, and the Swede score for the identification of HSIL+ in 525 women in Shanghai, China, referred for colposcopy with suspicious-looking cervixes (including cervixes with abnormal bleeding or obvious contact bleeding, abnormal vaginal discharge, recurrent erosion, cervical polyp, leukoplakia, condyloma, gross neoplasm, irregular surface, or cervical canal stenosis, or barrel-like cervixes), abnormal cervical cytology (ASC-US+), or positive hrHPV test results. The results showed that the colposcopic accuracy was lower with the RCI and the Swede score than with the IFCPC classification; the sensitivity of the RCI for identification of HSIL+ was 38% and the specificity was 95%, and the sensitivity of the Swede score for identification of HSIL+ was 13% and the specificity was 99%; these scores are currently not widely used. For the IFCPC classification, the sensitivity for identification of HSIL+ was estimated to be 64% and the specificity 96%. However, no unique classification has yet been adopted in clinical practice worldwide.

(d) Colposcopy training

Expertise in performing colposcopic examinations is attained and maintained by comprehensive training and experience with an adequate caseload. However, colposcopy training and assessment is neither uniform nor quality-assured worldwide. Even within the same country, there is considerable variation among colposcopists in training and experience (Wright, 2017).

Scientific colposcopy societies recognize the need to develop colposcopy standards for quality, and some have recently published training programmes (Public Health England, 2016; Mayeaux et al., 2017; Prendiville & Sankaranarayanan, 2017; AEPCC, 2018). Different societies propose different requirements, and few societies provide committees or infrastructures to support and oversee the training programmes (Moss et al., 2015). Nonetheless, most experts agree that training should involve supervised and unsupervised colposcopic assessment as well as attendance at clinical, histopathological, and cytopathological sessions (Public Health England, 2016; Prendiville, 2022).

Once a colposcopist is trained, performing a sufficient number of colposcopies per year is necessary to ensure continuing competence. The number differs between national colposcopy societies (Moss et al., 2013; Société Française de Colposcopie et de Pathologie Cervico-Vaginale, 2014; Public Health England, 2016; IFCPC, 2021), and some scientific groups do not specify the number of colposcopic evaluations needed per year to maintain competence (Mayeaux et al., 2017; Prendiville & Sankaranarayanan, 2017; AEPCC, 2018).

The systematic review by Mayeaux et al. (2017) of the different international guidelines for colposcopy quality described the wide variation between colposcopy societies in both colposcopy guidance and quality indicators, and emphasized the need for the standardization of guidance.

4.5.2 Accuracy of colposcopy in cytology-based screening

Despite the central role of colposcopy and colposcopy-directed biopsy in detecting cervical HSIL (Darragh et al., 2012), most of the available studies have evaluated colposcopy to assess the risk of underlying precancer or cancer. A limited number of studies have presented specific data for HSIL/CIN3+. However, recent studies evaluating colposcopy have shown that risk estimates for HSIL/CIN3+ were much less heterogeneous than results for HSIL/CIN2+; this probably reflects

the known variability and lack of reproducibility of CIN2/CIN3 diagnoses (Carreon et al., 2007; Herbert et al., 2008).

Four systematic reviews or meta-analyses have been performed on the accuracy of diagnostic colposcopy applied to women referred with abnormal cytology (Mitchell et al., 1998; Olaniyan, 2002; Mustafa et al., 2016; Brown & Tidy, 2019) (Table 4.36; web only; available from https://publications.iarc.fr/604). The most recent meta-analysis (Brown & Tidy, 2019), which included 10 973 women referred for colposcopy after abnormal cytology, reported a weighted mean sensitivity for histologically verified CIN2+ at a threshold of "any colposcopic abnormality" of 96% (range, 83–100%) and a weighted mean specificity of 34% (range, 5–67%). At a threshold of "high-grade colposcopic impression", the pooled sensitivity was 68% (range, 30–95%) and the pooled specificity was 76% (range, 48–97%). [The methods used for the calculation of diagnostic accuracy in clinical colposcopy trials are subject to several types of bias. The use of punch biopsies as the reference standard has been questioned in comparison with the results from excisional treatment after punch biopsy. It is important to consider that in many clinics biopsy is performed only when there is suspicion of disease. As a result, verification by biopsy is performed only when the outcome of colposcopy is positive and not when the outcome is negative. This form of bias results in overestimation of the sensitivity and underestimation of the specificity (Walter, 1999).]

Some analyses have attempted to eliminate this risk of bias. Underwood et al. (2012), in their systematic review, compared 7873 cases of colposcopy-directed cervical punch biopsy with their paired definitive histology from an excisional cervical biopsy or hysterectomy. At a threshold of "any colposcopic abnormality", the pooled sensitivity for a punch biopsy performed to diagnose a CIN2+ present in the surgical specimen was 91% (95% CI, 85–95%) and the pooled specificity was 25% (95% CI, 16–36%). At a threshold of "high-grade colposcopic impression", the pooled sensitivity was 80% (95% CI, 73–86%) and the pooled specificity was 63% (95% CI, 51–77%). Three subsequent retrospective studies (Kahramanoglu et al., 2019; Stuebs et al., 2019; Kim et al., 2020) evaluated the accuracy of colposcopy-directed biopsies with a paired specimen from an excisional treatment (including hysterectomy) and reported a sensitivity of punch biopsy for HSIL+/CIN2+ of 88–90% (92% in women with the entire TZ visible) and variable specificity of 37–59%. [None of these three studies specified whether the biopsies were performed for any colposcopy abnormality or only if a high-grade lesion was suspected.]

4.5.3 Colposcopy in HPV-based screening

When a transition is made from cytology-based strategies to strategies based on HPV testing, the central diagnostic role of colposcopy is maintained but the clinical characteristics of the patients and the number of women referred for colposcopy change profoundly. A major concern with switching from cytology to primary HPV screening is the management of HPV-positive women.

A study in 8369 women in the Guanacaste cohort study in Costa Rica (Porras et al., 2012) compared colposcopy characteristics and performance in women referred for colposcopy based on conventional cytology-based screening (ASC-US+) versus women with positive results in HPV-based screening (HPV typing using type-specific probes). The absolute risks of histological CIN2+ in women with abnormal colposcopy (or PPV) after cytology-based or HPV-based screening were similar (47.8% vs 41.5%, respectively; $P = 0.15$ for women aged 30 years or older). Similarly, there was no difference when ruling out histological CIN2+ in women with normal colposcopy (or NPV) in a cytology-based compared with an HPV-based

screening programme (87.2% vs 87.0%; $P = 0.92$ in women aged 30 years or older).

Colposcopy referrals for HPV-based screening compared with cytology-based screening were discussed in Section 4.4.2. To avoid overburdening the health-care system and overtreating women who are at low risk, a risk-based approach is needed to manage women with a positive HPV screening test result. A triage strategy enables the identification of HPV-positive women who are at higher risk of HSIL+ and who would most benefit from colposcopic examination. The different triage strategies were analysed in Section 4.4.7.

4.5.4 Random biopsies for diagnosis of CIN2+

In cervical cancer screening, it is especially important to rule out HSIL/CIN3+ in women with normal colposcopy, because most of these women do not undergo biopsy but are followed up.

In the Shanxi Province Cervical Cancer Screening Study I (SPOCCS I), Pretorius et al. (2004) evaluated colposcopies of 364 women in Shanxi Province, China, who were referred for colposcopy after an abnormal screening test with an entirely visible TZ in which all colposcopically abnormal areas were biopsied. If the colposcopic examination showed no lesion in a quadrant, a non-directed (random) biopsy was obtained within the TZ in that quadrant. In addition, endocervical curettage was performed after the cervical biopsies. The diagnosis of CIN2+ was made on a colposcopy-directed biopsy in 57% of women, a random biopsy in 37% of women, and an endocervical curettage in 6% of women.

Bekkers et al. (2008) evaluated the accuracy of colposcopy for the identification of HSIL in 6020 women in Melbourne, Australia, for whom the colposcopic impression was correlated with the histopathology result. In this study, colposcopy had a sensitivity of 60% and a PPV of 60% for the identification of HSIL, and the colposcopy-directed biopsies missed 39% of the HSIL. The sensitivity of colposcopy for the identification of HSIL was significantly higher ($P < 0.001$) with junior colposcopists (66.7%) than with senior colposcopists (57.5%), but the PPV was significantly lower ($P < 0.001$) with junior colposcopists (56%) than with senior colposcopists (64%).

In the analysis of the two studies in Shanxi Province, China (SPOCCS I and II), which evaluated 1383 women with abnormal cytology who were referred for colposcopy (Pretorius et al., 2011), 25% of the 222 CIN3+ and 10% of the 31 cervical cancers were diagnosed in a random biopsy. [The sensitivity of colposcopy for diagnosis of CIN3+ varied significantly among the seven physicians performing colposcopy, from 29% to 93% ($P < 0.001$).]

Other studies did not report a benefit from random biopsies. In the Evaluating the Visual Appearance of Cervical Lesions in Relation to its Histological Diagnosis, Human Papillomavirus Genotype and Other Viral Parameters (EVAH) study in the Netherlands and Spain, van der Marel et al. (2014) evaluated the benefit of random biopsies performed in 610 women referred for colposcopy after an abnormal cytology result. Multiple directed biopsies were collected from lesions, and a non-directed biopsy of normal-appearing tissue was added if fewer than four biopsies were collected. In women with at least two lesion-directed biopsies, the yield for CIN2+ increased from 51.7% (95% CI, 45.7–57.7%) for one directed biopsy to 60.4% (95% CI, 54.4–66.2%; $P < 0.001$) for two biopsies. An additional 5% of CIN2+ were detected in biopsies from women who had been underdiagnosed by colposcopy.

In the Biopsy Study of the University of Oklahoma Health Sciences Center and the United States National Cancer Institute (Wentzensen et al., 2015), only 2% of all HSIL diagnosed in the 690 participants were detected by random biopsies performed on a normal-appearing TZ.

A retrospective follow-up study in the setting of the National Health Service (NHS) Cervical Screening Programme in England within the HPV or LBC pilot studies (Kelly et al., 2012) evaluated the risk of incident CIN2+ in 1063 HPV-positive women with low-grade cytological abnormalities (ASC-US or LSIL) who had a normal colposcopy with a completely visible TZ. In these women, the cumulative rate of CIN2+ at 3 years of follow-up was 4.4% (95% CI, 4–7%), independent of the age of the woman.

In the TOMBOLA trial, 884 women aged 20–59 years, with the same inclusion criteria as in the study of Kelly et al. (2012), were evaluated to determine the rate of CIN2+ over 3 years of cervical cytology follow-up including an exit colposcopic examination (Cruickshank et al., 2015). CIN2+ was detected in 5% of the women at the end of the study.

Munmany et al. (2018) evaluated the accuracy of colposcopic evaluation at the time of large loop excision of the transformation zone (LLETZ), also known as loop electrosurgical excision procedure (LEEP), to identify women with a previous biopsy diagnosis of HSIL/CIN2/3 with a low probability of dysplasia at the time of treatment. Of 162 women included in the study, 34 (21%) had a normal colposcopy with a completely visible TZ, and the absence of LSIL (CIN1) or HSIL/CIN2/3 in the excised specimen was confirmed in 28 (82%) of the 34 women.

Overall, these studies indicate that in countries in which colposcopy is part of a properly constructed, quality-assured programme, a normal colposcopy is associated with a very high NPV.

4.5.5 Risk-based colposcopy practice

Women referred for colposcopy after an abnormal screening result have a wide range of risk of harbouring a cervical lesion. Recently, it has been suggested that the risk of underlying histological HSIL can be estimated before colposcopic evaluation by assessing the information provided by the screening test (cytology and/or molecular test results). In this strategy, the practice of colposcopy and biopsy can be modified depending on the risk of precancer (Wentzensen et al., 2017; AEPCC, 2018; Perkins, et al., 2020). The risk of cervical precancer can be based on the results of the screening and follow-up tests (Dillner et al., 2008; Schiffman et al., 2015; Castle et al., 2016; Wentzensen et al., 2017; AEPCC, 2018; de Sanjosé et al., 2018; Egemen et al., 2020; Perkins, et al., 2020), as summarized in Table 4.37 (web only; available from https://publications.iarc.fr/604).

Moreover, information provided by the colposcopic impression may modify the need to perform multiple biopsies, including random biopsies (Wentzensen & Clarke, 2017; AEPCC, 2018; Silver et al., 2018; Egemen et al., 2020). A recent meta-analysis evaluated the risk strata based on combinations of cytology, HPV16 and/or HPV18 genotyping, and colposcopic impression (Silver et al., 2018). Eligible studies reported colposcopic impression and either cytology results or HPV16/18 partial genotyping results as well as a histological biopsy diagnosis from adult women. Women with < HSIL cytology who were HPV16/18-negative and had a normal colposcopic impression had the lowest risk of prevalent precancer and cancer (< 0.5% for HSIL/CIN3+). Women with at least two of the three high-risk results (i.e. HSIL cytology, HPV16- and/or HPV18-positive, and grade 2 changes at colposcopy) were at high risk (29–53% for HSIL/CIN3+), and women with all three of these high-risk results had the highest risk (> 70% for HSIL/CIN3+). Table 4.38 shows the levels (low, intermediate, and high) of risk of histological HSIL on the basis of cytology, HPV testing, and colposcopic findings.

On the basis of the current evidence, scientific societies have issued new colposcopy standards and risk-based management guidelines for the low-risk and high-risk groups of women based on

Table 4.38 Levels of risk of histological HSIL on the basis of cytology, HPV testing, and colposcopic findings

Low risk	Intermediate risk	High risk
Fulfil the following 3 criteria: • Cytology < HSIL • No HPV16/18 • Normal colposcopy	Cases not included in the other 2 risk groups	Fulfil at least 2 of the following 3 criteria: • Cytology ≥ HSIL, AGC, or ASC-H • HPV16 and/or HPV18 • Colposcopy showing grade 2 changes (high-grade/HSIL)

AGC, atypical glandular cells; ASC-H, atypical squamous cells, cannot exclude high-grade squamous epithelial lesions; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion.
Reproduced with permission from AEPCC (2018).

the available test results (cytology, HPV testing, and colposcopic impression) (Wentzensen et al., 2017; AEPCC, 2018; Perkins et al., 2020). Random biopsies should not be performed for women with < HSIL cytology who are HPV16/18-negative and have normal colposcopy. In contrast, in the case of abnormal colposcopy, even without any suspicion of cervical HSIL, cervical biopsy should be performed in women with HSIL cytology and/or HPV16- and/or HPV18-positive tests, particularly where adequate training and quality assurance are not in place. In women in the highest-risk group, the benefit of taking random biopsies from normal colposcopic areas within the TZ could also be considered. When multiple biopsies are taken and are negative, it is mandatory to provide close follow-up of the woman (i.e. every 6 months) (AEPCC, 2015), and if high-grade abnormalities (HSIL cytology and/or colposcopy showing grade 2 changes with negative biopsies) persist in the follow-up tests, type 3 excision (Bornstein et al., 2012) should be considered (Del Pino et al., 2010; AEPCC, 2015, 2018). In contrast, expedited excisional treatment (defined as excisional treatment without preceding colposcopy-directed biopsy demonstrating histological HSIL/CIN2+) is entirely appropriate in selected women at very high risk of harbouring HSIL/CIN3+, according to clinical guidelines (Wentzensen et al., 2017; Wright, 2017; Egemen et al., 2020; Perkins et al., 2020) (see also Section 1.2.5).

The main advantage of risk stratification is that the colposcopic examination and the biopsy strategy are adapted to the risk stratum. The colposcopist can either not perform a biopsy (in women at low risk) or perform expedited excisional treatment (in women at high risk). In women at intermediate risk, colposcopy-directed biopsies are appropriate. The potential benefit of biopsies in minimal acetowhite areas or when the colposcopy is normal (random biopsies) should be considered in each case (Waxman et al., 2017; Wentzensen et al., 2017; AEPCC, 2018).

4.5.6 Harmful effects of colposcopy

The harmful effects of colposcopy are (i) harms related to the procedure, (ii) harms linked with inadequate indication for colposcopy, and (iii) harms related to lack of experience or quality assurance.

(a) Harms related to the procedure

(i) Pain or discomfort

Although colposcopy is generally a well-tolerated examination, and therefore administration of analgesic drugs before the procedure is not recommended, some women may report discomfort due to the prolonged placing of the speculum or the application of acetic acid or iodine solution, or cramping or pain associated with the biopsy procedure (Khan et al., 2017;

AEPCC, 2018). In the TOMBOLA trial (Sharp et al., 2009), of the 401 women who underwent colposcopic examination (without biopsy or treatment), 18% (95% CI, 15–23%) reported some pain or physical discomfort when questioned at 6 weeks and 4 months after a colposcopy, and 5% (95% CI, 3–8%) reported that the discomfort was moderate to severe. O'Connor et al. (2017) reported that 59% of 248 women questioned at 4, 8, and 12 months after a colposcopy described pain (75% of the procedures included punch biopsies or conization). Pain during colposcopy is more closely related to the biopsy procedure or the treatment than to the colposcopy procedure itself. In addition, in the TOMBOLA trial (Sharp et al., 2009), of the women who underwent colposcopic examination (without biopsy or treatment), 18% (95% CI, 15–23%) reported pain; this proportion increased to 53% (95% CI, 44–61%) for those who underwent colposcopy and punch biopsy and to 67% (95% CI, 59–74%) for those who underwent colposcopy and excisional treatment (conization).

Pain and discomfort are generally experienced at the time of the procedure, but sometimes cramping can persist for a few hours. On the basis of two RCTs including 129 women, a Cochrane review concluded that there was no difference in pain relief between women undergoing colposcopy (without treatment) who received oral analgesics and those who received placebo or no treatment (mean difference, −3.51; 95% CI, −10.03 to 3.01 [low-quality evidence]) (Gajjar et al., 2016).

A prospective study conducted at Concord Women's Health Center in Israel including 101 women who underwent colposcopy reported a negative correlation between age and pain associated with the procedure (Pearson correlation coefficient, −0.220; $P < 0.05$) (Handelzalts et al., 2015).

(ii) Anxiety

Anxiety, worry, and fear are the feelings most commonly described during colposcopy (Galaal et al., 2011; O'Connor et al., 2016). In a systematic review evaluating psychological outcomes after colposcopy and related procedures, which included 16 studies (O'Connor et al., 2016), 60% of women undergoing colposcopy for the first time experienced anxiety (defined as an STAI score > 35), and 18% reported high anxiety levels (defined as an STAI score > 44); also, one third of the women undergoing colposcopy for the first time experienced distress or worry. The results of the procedure had impacts on the course of the negative feelings. At 6 weeks after the procedure, 21% of the women with a normal TZ and 42% of the women with an abnormal TZ still had significant distress. Moreover, in women with a normal TZ, distress and worry were significantly increased in those who reported pain or discharge after the procedure (Sharp et al., 2011, 2013).

Many women also report worry or anxiety in the period between the time of being notified of an abnormal screening result and the colposcopy appointment (Khan et al., 2017; Young et al., 2018). although it is unclear whether the diagnosis of an abnormal screening test or the colposcopy itself contributes to negative feelings (Khan et al., 2017). In general, women are less concerned about the procedure itself and are more anxious about having an HPV infection or cancer (see Section 4.4.8). Waller et al. (2007) evaluated the psychosocial impact of having a second positive HPV test result in 30 women undergoing cervical cancer screening who were HPV-positive with normal cytology at the first visit, and who attended for a repeat HPV test 12 months later. The study found that women appeared to be more distressed by a second positive HPV test result than by the first one. They also expressed a clear preference for immediate colposcopy over continued surveillance,

indicating that the anxiety was associated mainly with the screening result but also with a desire for a speedy resolution and fears about progression to cancer.

Colposcopy may also have a negative influence on sexual function. Seven studies included in the systematic review by O'Connor et al. (2016) assessed some aspect of sexual or psychosexual functioning after colposcopy. Although one study reported that the mean total score in the Female Sexual Function Index (FSFI) after colposcopy was above the threshold for female sexual disorder, the other studies comparing pre- with post-colposcopy sexual or psychosexual functioning reported conflicting results, with no consistent pattern of impact. [This secondary effect may be more closely related to abnormal screening test results than to the colposcopy procedure itself.]

Different approaches have been evaluated to reduce anxiety in women undergoing colposcopy after an abnormal screening test. Effective information and communication have consistently been shown to reduce anxiety (Kola et al., 2013; Handelzalts et al., 2015). Women who have not been extensively informed and are unaware of the possibility of experiencing side-effects score significantly higher for distress and anxiety during follow-up (O'Connor et al., 2017). Video colposcopy, which enables women to observe their own anatomy and watch what the colposcopist is doing, has been reported to reduce anxiety, in some studies (Kola et al., 2013) but not in others (Hilal et al., 2017).

Music therapy has been used to reduce anxiety associated with various medical procedures; however, in a recent meta-analysis, music therapy had no positive effect on reducing anxiety or pain or increasing satisfaction levels during colposcopy (Abdelhakim et al., 2019).

Most studies on the psychological impact of colposcopy have been performed in women undergoing colposcopy for the first time. However, compared with women undergoing subsequent colposcopic examinations, those undergoing colposcopy for the first time typically experience increased anxiety both before and after colposcopy and display a tendency to seek information about the procedure (Handelzalts et al., 2015).

(iii) Anaphylactic reaction to iodine solution

Isolated examples of allergic reactions to iodine solution have been described. These include pruritus, vaginal oedema, hypotension, tachycardia, and breathing difficulties. The symptoms usually disappear upon withdrawal of the iodine solution (Indraccolo et al., 2009).

(b) Harms linked with inadequate indication for colposcopy

Although colposcopy was initially used as a tool for primary screening of cervical cancer and precancer, an increased understanding of the natural history of HPV infection and its progression to cervical neoplasia has recently reduced the indications for colposcopy. Strict adherence to indications for colposcopy (Table 4.33) minimizes the side-effects associated with inappropriate use of this procedure.

(c) Harms related to lack of experience or quality assurance

Colposcopy requires adequate training and experience to attain proficiency, assure quality, and maintain competence in performing the procedure. The proportion of false-negative results of colposcopy (women with HSIL/CIN2+ classified as being disease-free) correlates directly with the expertise of the colposcopist.

As mentioned above, one study showed significantly higher sensitivity for the identification of HSIL when performed by junior colposcopists (with 0–2 years of experience in colposcopy) compared with senior colposcopists (with > 3 years of experience) (66.7% vs 57.5%; $P < 0.001$), but a significantly lower PPV (56% vs 64%; $P < 0.001$) (Bekkers et al., 2008).

A retrospective analysis comparing the precision of diagnosis by colposcopy-directed biopsy with the final histological outcome of the surgical specimen in 641 women showed a risk of underdiagnosis of HSIL (false negativity) of 12% when the colposcopist had 0–5 years of experience and of 8% when the colposcopist had more than 10 years of experience (Stuebs et al., 2019).

References

Abdelhakim AM, Samy A, Abbas AM (2019). Effect of music in reducing patient anxiety during colposcopy: a systematic review and meta-analysis of randomized controlled trials. *J Gynecol Obstet Hum Reprod*. 48(10):855–61. doi:10.1016/j.jogoh.2019.07.007 PMID:31276848

AEPCC (2015). Guía: prevención del cáncer de cuello de útero. Asociación Española de Patología Cervical y Colposcopia. [in Spanish] Available from: http://www.aepcc.org/wp-content/uploads/2016/01/AEPCC_revista02.pdf.

AEPCC (2018). Guía de colposcopia. Estándares de calidad. Asociación Española de Patología Cervical y Colposcopia. [in Spanish] Available from: http://www.aepcc.org/wp-content/uploads/2019/01/AEPCC_revista10-colposcopia-web.pdf.

Alan M, Gunyeli I, Gultekin M, Sancı M, Yuce K (2020). Correlation of Swede score colposcopy scoring system and histopathological results in patients with high-risk HPV infection other than HPV16 and 18. *Int J Gynecol Cancer*. 30(1):35–40. doi:10.1136/ijgc-2019-000932 PMID:31792083

Bappa LA, Yakasai IA (2013). Colposcopy: the scientific basis. *Ann Afr Med*. 12(2):86–9. doi:10.4103/1596-3519.112396 PMID:23713014

Bekkers RL, van de Nieuwenhof HP, Neesham DE, Hendriks JH, Tan J, Quinn MA (2008). Does experience in colposcopy improve identification of high grade abnormalities? *Eur J Obstet Gynecol Reprod Biol*. 141(1):75–8. doi:10.1016/j.ejogrb.2008.07.007 PMID:18760872

Beller FK, Khatamee M (1966). Evaluation of punch biopsy of the cervix under direct colposcopic observation (target punch biopsy). *Obstet Gynecol*. 28(5):622–5. PMID:5923541

Beyer DA, Rody A, Cirkel C, Schmidt N, Neumann K (2017). Mandatory colposcopic findings of severe cervical dysplasia. Are there key-signs that need our special attention? *J Gynecol Obstet Hum Reprod*. 46(8):643–6. doi:10.1016/j.jogoh.2017.07.002 PMID:28698072

Bornstein J, Bentley J, Bösze P, Girardi F, Haefner H, Menton M, et al. (2012). 2011 colposcopic terminology of the International Federation for Cervical Pathology and Colposcopy. *Obstet Gynecol*. 120(1):166–72. doi:10.1097/AOG.0b013e318254f90c PMID:22914406

Brown BH, Tidy JA (2019). The diagnostic accuracy of colposcopy – a review of research methodology and impact on the outcomes of quality assurance. *Eur J Obstet Gynecol Reprod Biol*. 240:182–6. doi:10.1016/j.ejogrb.2019.07.003 PMID:31302386

Carreon JD, Sherman ME, Guillén D, Solomon D, Herrero R, Jerónimo J, et al. (2007). CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. *Int J Gynecol Pathol*. 26(4):441–6. doi:10.1097/pgp.0b013e31805152ab PMID:17885496

Castle PE, Aslam S, Behrens C (2016). Cervical precancer and cancer risk by human papillomavirus status and cytologic interpretation: implications for risk-based management. *Cancer Epidemiol Biomarkers Prev*. 25(12):1595–9. doi:10.1158/1055-9965.EPI-16-0330 PMID:27587789

Colgan TJ, Lickrish GM (1990). The topography and invasive potential of cervical adenocarcinoma *in situ*, with and without associated squamous dysplasia. *Gynecol Oncol*. 36(2):246–9. doi:10.1016/0090-8258(90)90182-K PMID:2298414

Cruickshank ME, Cotton SC, Sharp L, Smart L, Walker LG, Little J; TOMBOLA Group (2015). Management of women with low grade cytology: how reassuring is a normal colposcopy examination? *BJOG*. 122(3):380–6. doi:10.1111/1471-0528.12932 PMID:24947656

Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, et al.; Members of LAST Project Work Groups (2012). The Lower Anogenital Squamous Terminology Standardization Project for HPV-Associated Lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. [Erratum in: Arch Pathol Lab Med. 2013 137(6):738] *Arch Pathol Lab Med*. 136(10):1266–97. doi:10.5858/arpa.LGT200570 PMID:22742517

de Sanjosé S, Rodríguez-Salés V, Bosch XF, Ibañez R, Bruni L (2018). Population-based e-records to evaluate HPV triage of screen-detected atypical squamous cervical lesions in Catalonia, Spain, 2010–15. *PLoS One*. 13(11):e0207812. doi:10.1371/journal.pone.0207812 PMID:30475876

Del Pino M, Torne A, Alonso I, Mula R, Masoller N, Fuste V, et al. (2010). Colposcopy prediction of progression in human papillomavirus infections with minor cervical lesions. *Obstet Gynecol*. 116(6):1324–31. doi:10.1097/AOG.0b013e3181f74885 PMID:21099598

Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. (2008). Long term predictive values of cytology and human papillomavirus testing in cervical

cancer screening: joint European cohort study. *BMJ*. 337:a1754. doi:10.1136/bmj.a1754 PMID:18852164

Egemen D, Cheung LC, Chen X, Demarco M, Perkins RB, Kinney W, et al. (2020). Risk estimates supporting the 2019 ASCCP Risk-based Management Consensus Guidelines. *J Low Genit Tract Dis*. 24(2):132–43. doi:10.1097/LGT.0000000000000529 PMID:32243308

Ferris DG, Litaker M; ALTS Group (2005). Interobserver agreement for colposcopy quality control using digitized colposcopic images during the ALTS trial. *J Low Genit Tract Dis*. 9(1):29–35. doi:10.1097/00128360-200501000-00007 PMID:15870519

Gajjar K, Martin-Hirsch PPL, Bryant A, Owens GL (2016). Pain relief for women with cervical intraepithelial neoplasia undergoing colposcopy treatment. *Cochrane Database Syst Rev*. 7(7):CD006120. doi:10.1002/14651858.CD006120.pub4 PMID:27428114

Galaal K, Bryant A, Deane KHO, Al-Khaduri M, Lopes AD (2011). Interventions for reducing anxiety in women undergoing colposcopy. *Cochrane Database Syst Rev*. (12):CD006013. doi:10.1002/14651858.CD006013.pub3 PMID:22161395

Handelzalts JE, Krissi H, Levy S, Broitman M, Binyamin L, Peled Y (2015). Multidimensional associations of pain and anxiety before and after colposcopy. *Int J Gynaecol Obstet*. 131(3):297–300. doi:10.1016/j.ijgo.2015.05.038 PMID:26386494

Herbert A, Arbyn M, Bergeron C (2008). Why CIN3 and CIN2 should be distinguished on histological reports. *Cytopathology*. 19(1):63–4. doi:10.1111/j.1365-2303.2006.00539.x PMID:18205632

Hilal Z, Alici F, Tempfer CB, Seebacher V, Rezniczek GA (2017). Video colposcopy for reducing patient anxiety during colposcopy: a randomized controlled trial. *Obstet Gynecol*. 130(2):411–9. doi:10.1097/AOG.0000000000002127 PMID:28697117

Hilgarth M, Menton M (1996). The colposcopic screening. *Eur J Obstet Gynecol Reprod Biol*. 65(1):65–9. doi:10.1016/0028-2243(95)02306-D PMID:8706960

Hinselmann H (1925). Verbesserung der Inspektionsmöglichkeiten von Vulva, Vagina und Portio. *Munch Med Wochenschr*. 72:1733.

IFCPC (2021). IFCPC-IARC training course in colposcopy and the prevention of cervical cancer. Rockwell (MD), USA: International Federation of Cervical Pathology and Colposcopy. Available from: https://ifcpc.org/education.

Indraccolo U, Palombino K, Greco P (2009). Anaphylactic-like reaction to Lugol solution during colposcopy. *South Med J*. 102(1):96–7. doi:10.1097/SMJ.0b013e31818894f4 PMID:19077773

Jordan JA (1985). Colposcopy in the diagnosis of cervical cancer and precancer. *Clin Obstet Gynaecol*. 12(1):67–76. PMID:3995857

Kahramanoglu I, Demirkiran F, Turan H, Bese T, Yilmaz N, Ilvan S, et al. (2019). The use of colposcopic punch biopsy in the management of abnormal cervical cytology: a 5-year retrospective audit. *J Obstet Gynaecol*. 39(1):110–4. doi:10.1080/01443615.2018.1468740 PMID:30230394

Kelly RS, Walker P, Kitchener H, Moss SM (2012). Incidence of cervical intraepithelial neoplasia grade 2 or worse in colposcopy-negative/human papillomavirus-positive women with low-grade cytological abnormalities. *BJOG*. 119(1):20–5. doi:10.1111/j.1471-0528.2011.02970.x PMID:21624034

Khan MJ, Werner CL, Darragh TM, Guido RS, Mathews C, Moscicki AB, et al. (2017). ASCCP Colposcopy Standards: role of colposcopy, benefits, potential harms, and terminology for colposcopic practice. *J Low Genit Tract Dis*. 21(4):223–9. doi:10.1097/LGT.0000000000000338 PMID:28953110

Kim SI, Kim SJ, Suh DH, Kim K, No JH, Kim YB (2020). Pathologic discrepancies between colposcopy-directed biopsy and loop electrosurgical excision procedure of the uterine cervix in women with cytologic high-grade squamous intraepithelial lesions. *J Gynecol Oncol*. 31(2):e13. doi:10.3802/jgo.2020.31.e13 PMID:31912671

Kola S, Walsh JC, Hughes BM, Howard S (2013). Matching intra-procedural information with coping style reduces psychophysiological arousal in women undergoing colposcopy. *J Behav Med*. 36(4):401–12. doi:10.1007/s10865-012-9435-z PMID:22729966

Leeson SC, Alibegashvili T, Arbyn M, Bergeron C, Carriero C, Mergui JL, et al. (2014). The future role for colposcopy in Europe. *J Low Genit Tract Dis*. 18(1):70–8. doi:10.1097/LGT.0b013e318286b899 PMID:23774077

Li Y, Duan X, Sui L, Xu F, Xu S, Zhang H, et al. (2017). Closer to a uniform language in colposcopy: study on the potential application of 2011 International Federation for Cervical Pathology and Colposcopy terminology in clinical practice. *Biomed Res Int*. 2017:8984516. doi:10.1155/2017/8984516 PMID:28626767

Mayeaux EJ, Cox JT, editors (2013). Colposcopia: texto y atlas. 3rd ed. Barcelona, Spain: Lippincott Williams & Wilkins.

Mayeaux EJ Jr, Novetsky AP, Chelmow D, Choma K, Garcia F, Liu AH, et al. (2017). Systematic review of international colposcopy quality improvement guidelines. *J Low Genit Tract Dis*. 21(4):249–57. doi:10.1097/LGT.0000000000000344 PMID:28953114

Mitchell MF, Schottenfeld D, Tortolero-Luna G, Cantor SB, Richards-Kortum R (1998). Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. *Obstet Gynecol*. 91(4):626–31. doi:10.1097/00006250-199804000-00029 PMID:9540955

Moss EL, Arbyn M, Dollery E, Leeson S, Petry KU, Nieminen P, et al. (2013). European Federation of Colposcopy quality standards Delphi consultation. *Eur J Obstet Gynecol Reprod Biol*. 170(1):255–8. doi:10.1016/j.ejogrb.2013.06.032 PMID:23891389

Moss EL, Redman CWE, Arbyn M, Dollery E, Petry KU, Nieminen P, et al. (2015). Colposcopy training and assessment across the member countries of the European Federation for Colposcopy. *Eur J Obstet Gynecol Reprod Biol*. 188:124–8. doi:10.1016/j.ejogrb.2015.03.012 PMID:25839437

Munmany M, Torné A, Nonell R, Barnadas E, Luqui N, Ordi J, et al. (2018). Colposcopy evaluation at the time of loop electrosurgical excision procedure may avoid unnecessary treatment. *J Low Genit Tract Dis*. 22(4):367–74. doi:10.1097/LGT.0000000000000410 PMID:29957657

Mustafa RA, Santesso N, Khatib R, Mustafa AA, Wiercioch W, Kehar R, et al. (2016). Systematic reviews and meta-analyses of the accuracy of HPV tests, visual inspection with acetic acid, cytology, and colposcopy. *Int J Gynaecol Obstet*. 132(3):259–65. doi:10.1016/j.ijgo.2015.07.024 PMID:26851054

O'Connor M, Gallagher P, Waller J, Martin CM, O'Leary JJ, Sharp L; Irish Cervical Screening Research Consortium (CERVIVA) (2016). Adverse psychological outcomes following colposcopy and related procedures: a systematic review. *BJOG*. 123(1):24–38. doi:10.1111/1471-0528.13462 PMID:26099164

O'Connor M, O'Brien K, Waller J, Gallagher P, D'Arcy T, Flannelly G, et al.; Irish Cervical Screening Research Consortium (CERVIVA) (2017). Physical after-effects of colposcopy and related procedures, and their inter-relationship with psychological distress: a longitudinal survey. *BJOG*. 124(9):1402–10.doi:10.1111/1471-0528.14671 PMID:28374937

Olaniyan OB (2002). Validity of colposcopy in the diagnosis of early cervical neoplasia – a review. *Afr J Reprod Health*. 6(3):59–69. doi:10.2307/3583258 PMID:12685410

Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al.; 2019 ASCCP Risk-Based Management Consensus Guidelines Committee (2020). 2019 ASCCP Risk-Based Management Consensus Guidelines for abnormal cervical cancer screening tests and cancer precursors. [Erratum in: J Low Genit Tract Dis. 2020 24(4):427] *J Low Genit Tract Dis*. 24(2):102–31. doi:10.1097/LGT.0000000000000525 PMID:32243307

Porras C, Wentzensen N, Rodríguez AC, Morales J, Burk RD, Alfaro M, et al. (2012). Switch from cytology-based to human papillomavirus test-based cervical screening: implications for colposcopy. *Int J Cancer*. 130(8):1879–87. doi:10.1002/ijc.26194 PMID:21607948

Prendiville W (2022). A systematic approach to colposcopic examination. In: Prendiville W, del Pino M, Wentzensen N, editors. Core topics in cervical dysplasia and colposcopy. Cambridge, UK: Cambridge University Press. (forthcoming)

Prendiville W, Sankaranarayanan R (2017). Colposcopy and treatment of cervical precancer (IARC Technical Publications, No. 45). Lyon, France: International Agency for Research on Cancer. Available from: https://publications.iarc.fr/555.

Pretorius RG, Belinson JL, Burchette RJ, Hu S, Zhang X, Qiao YL (2011). Regardless of skill, performing more biopsies increases the sensitivity of colposcopy. *J Low Genit Tract Dis*. 15(3):180–8. doi:10.1097/LGT.0b013e3181fb4547 PMID:21436729

Pretorius RG, Zhang WH, Belinson JL, Huang MN, Wu LY, Zhang X, et al. (2004). Colposcopically directed biopsy, random cervical biopsy, and endocervical curettage in the diagnosis of cervical intraepithelial neoplasia II or worse. *Am J Obstet Gynecol*. 191(2):430–4. doi:10.1016/j.ajog.2004.02.065 PMID:15343217

Public Health England (2016). NHS Cervical Screening Programme: colposcopy and programme management. 3rd ed. London, UK: Public Health England. Available from: https://www.gov.uk/government/publications/cervical-screening-programme-and-colposcopy-management.

Ranga R, Rai S, Kumari A, Mathur S, Kriplani A, Mahey R, et al. (2017). A comparison of the strength of association of Reid Colposcopic Index and Swede score with cervical histology. *J Low Genit Tract Dis*. 21(1):55–8. doi:10.1097/LGT.0000000000000278 PMID:27851696

Reid R, Campion MJ (1989). Lesiones del cuello uterino relacionadas con papilomavirus humano: biología y características colposcópica. [in Spanish] *Clin Ostet Ginecol*. 1:151–73.

Reid R, Scalzi P (1985). Genital warts and cervical cancer. VII. An improved colposcopic index for differentiating benign papillomaviral infections from high-grade cervical intraepithelial neoplasia. *Am J Obstet Gynecol*. 153(6):611–8. doi:10.1016/S0002-9378(85)80244-1 PMID:2998190

Ruiz Moreno JA (2010). Pasado, presente y futuro de la colposcopia: perspectiva de un profesional en tracto genital inferior. [in Spanish] *Arch Méd Actual Trac Gen Inf*. 2(2):27–34. Available from: http://www.medigraphic.com/pdfs/archivostgi/tgi-2010/tgi102e.pdf.

Schiffman M, Vaughan LM, Raine-Bennett TR, Castle PE, Katki HA, Gage JC, et al. (2015). A study of HPV typing for the management of HPV-positive ASC-US cervical cytologic results. *Gynecol Oncol*. 138(3):573–8. doi:10.1016/j.ygyno.2015.06.040 PMID:26148763

Schiller W (1933). Early diagnosis of carcinoma of the cervix. *Surg Gynecol Obstet*. 56:210–22.

Schulmeyer CE, Stübs F, Gass P, Renner SK, Hartmann A, Strehl J, et al. (2020). Correlation between referral cytology and in-house colposcopy-guided cytology for detecting early cervical neoplasia. *Arch Gynecol*

Obstet. 301(1):263–71. doi:10.1007/s00404-019-05389-1 PMID:31811416

Sharp L, Cotton S, Carsin AE, Gray N, Thornton A, Cruickshank M, et al.; TOMBOLA Group (2013). Factors associated with psychological distress following colposcopy among women with low-grade abnormal cervical cytology: a prospective study within the Trial Of Management of Borderline and Other Low-grade Abnormal smears (TOMBOLA). *Psychooncology*. 22(2):368–80. doi:10.1002/pon.2097 PMID:22162138

Sharp L, Cotton S, Cochran C, Gray N, Little J, Neal K, et al.; TOMBOLA Group (2009). After-effects reported by women following colposcopy, cervical biopsies and LLETZ: results from the TOMBOLA trial. *BJOG*. 116(11):1506–14. doi:10.1111/j.1471-0528.2009.02263.x PMID:19583712

Sharp L, Cotton S, Gray N, Avis M, Russell I, Walker L, et al.; TOMBOLA Group (2011). Long-term psychosocial impact of alternative management policies in women with low-grade abnormal cervical cytology referred for colposcopy: a randomised controlled trial. *Br J Cancer*. 104(2):255–64. doi:10.1038/sj.bjc.6606042 PMID:21179033

Silver MI, Andrews J, Cooper CK, Gage JC, Gold MA, Khan MJ, et al. (2018). Risk of cervical intraepithelial neoplasia 2 or worse by cytology, human papillomavirus 16/18, and colposcopy impression: a systematic review and meta-analysis. *Obstet Gynecol*. 132(3):725–35. doi:10.1097/AOG.0000000000002812 PMID:30095780

Société Française de Colposcopie et de Pathologie Cervico-Vaginale (2014). Charte de qualité en colposcopie. Available from: http://www.societe-colposcopie.com/grand-public/charte-de-qualite-en-colposcopie.

Soutter WP (1993). A practical guide to colposcopy. New York (NY), USA: Oxford University Press.

Stafl A (1976). New nomenclature for colposcopy. Report of the committee on terminology. *Obstet Gynecol*. 48(1):123–4. PMID:945521

Stafl A, Wilbanks GD (1991). An international terminology of colposcopy: report of the Nomenclature Committee of the International Federation of Cervical Pathology and Colposcopy. *Obstet Gynecol*. 77(2):313–4. doi:10.1097/00006250-199102000-00032 PMID:1988900

Strander B, Ellström-Andersson A, Franzén S, Milsom I, Rådberg T (2005). The performance of a new scoring system for colposcopy in detecting high-grade dysplasia in the uterine cervix. *Acta Obstet Gynecol Scand*. 84(10):1013–7. doi:10.1111/j.0001-6349.2005.00895.x PMID:16167921

Stuebs FA, Schulmeyer CE, Mehlhorn G, Gass P, Kehl S, Renner SK, et al. (2019). Accuracy of colposcopy-directed biopsy in detecting early cervical neoplasia: a retrospective study. *Arch Gynecol Obstet*. 299(2):525–32. doi:10.1007/s00404-018-4953-8 PMID:30367250

Underwood M, Arbyn M, Parry-Smith W, De Bellis-Ayres S, Todd R, Redman CW, et al. (2012). Accuracy of colposcopy-directed punch biopsies: a systematic review and meta-analysis. *BJOG*. 119(11):1293–301. doi:10.1111/j.1471-0528.2012.03444.x PMID:22882742

van der Marel J, van Baars R, Rodriguez A, Quint WGV, van de Sandt MM, Berkhof J, et al. (2014). The increased detection of cervical intraepithelial neoplasia when using a second biopsy at colposcopy. *Gynecol Oncol*. 135(2):201–7. doi:10.1016/j.ygyno.2014.08.040 PMID:25204846

van Niekerk WA, Dunton CJ, Richart RM, Hilgarth M, Kato H, Kaufman RH, et al. (1998). Colposcopy, cervicography, speculoscopy and endoscopy: IAC Task Force summary. *Acta Cytol*. 42(1):33–49. doi:10.1159/000331533 PMID:9479322

Vercellino GF, Erdemoglu E, Chiantera V, Vasiljeva K, Drechsler I, Cichon G, et al. (2013). Validity of the colposcopic criteria inner border sign, ridge sign, and rag sign for detection of high-grade cervical intraepithelial neoplasia. *Obstet Gynecol*. 121(3):624–31. doi:10.1097/AOG.0b013e3182835831 PMID:23635627

Walker P, Dexeus S, De Palo G, Barrasso R, Campion M, Girardi F, et al.; Nomenclature Committee of the International Federation for Cervical Pathology and Colposcopy (2003). International terminology of colposcopy: an updated report from the International Federation for Cervical Pathology and Colposcopy. *Obstet Gynecol*. 101(1):175–7. doi:10.1016/S0029-7844(02)02581-4 PMID:12517664

Waller J, McCaffery K, Kitchener H, Nazroo J, Wardle J (2007). Women's experiences of repeated HPV testing in the context of cervical cancer screening: a qualitative study. *Psychooncology*. 16(3):196–204. doi:10.1002/pon.1053 PMID:16858719

Walter SD (1999). Estimation of test sensitivity and specificity when disease confirmation is limited to positive results. *Epidemiology*. 10(1):67–72. doi:10.1097/00001648-199901000-00012 PMID:9888282

Waxman AG, Conageski C, Silver MI, Tedeschi C, Stier EA, Apgar B, et al. (2017). ASCCP Colposcopy Standards: how do we perform colposcopy? Implications for establishing standards. *J Low Genit Tract Dis*. 21(4):235–41. doi:10.1097/LGT.0000000000000336 PMID:28953112

Wentzensen N, Clarke MA (2017). From clinical epidemiology to practice recommendations: knowledge gaps and uncertainty in the management of anal precancers. *Cancer*. 123(23):4530–4. doi:10.1002/cncr.31033 PMID:28949415

Wentzensen N, Schiffman M, Silver MI, Khan MJ, Perkins RB, Smith KM, et al. (2017). ASCCP Colposcopy Standards: risk-based colposcopy practice. *J Low Genit Tract Dis*. 21(4):230–4. doi:10.1097/LGT.0000000000000334 PMID:28953111

Wentzensen N, Walker JL, Gold MA, Smith KM, Zuna RE, Mathews C, et al. (2015). Multiple biopsies and detection of cervical cancer precursors at colposcopy. *J Clin Oncol*. 33(1):83–9. doi:10.1200/JCO.2014.55.9948 PMID:25422481

Wright TC Jr (2017). The new ASCCP Colposcopy Standards. *J Low Genit Tract Dis*. 21(4):215. doi:10.1097/LGT.0000000000000337 PMID:28953108

Young A, Cotton S, Cruickshank ME (2018). Information needs of young women vaccinated against HPV attending colposcopy: a qualitative study. *BMC Womens Health*. 18(1):200. doi:10.1186/s12905-018-0691-0 PMID:30541542

4.6 Emerging technologies

Recent advances in understanding of HPV-associated carcinogenesis have led to the development and evaluation of many new technologies and approaches for cervical cancer screening, triage, management, and diagnosis. Three types of approaches for the detection of cervical precancer are distinguished: those based on visual, cytological, and molecular technologies.

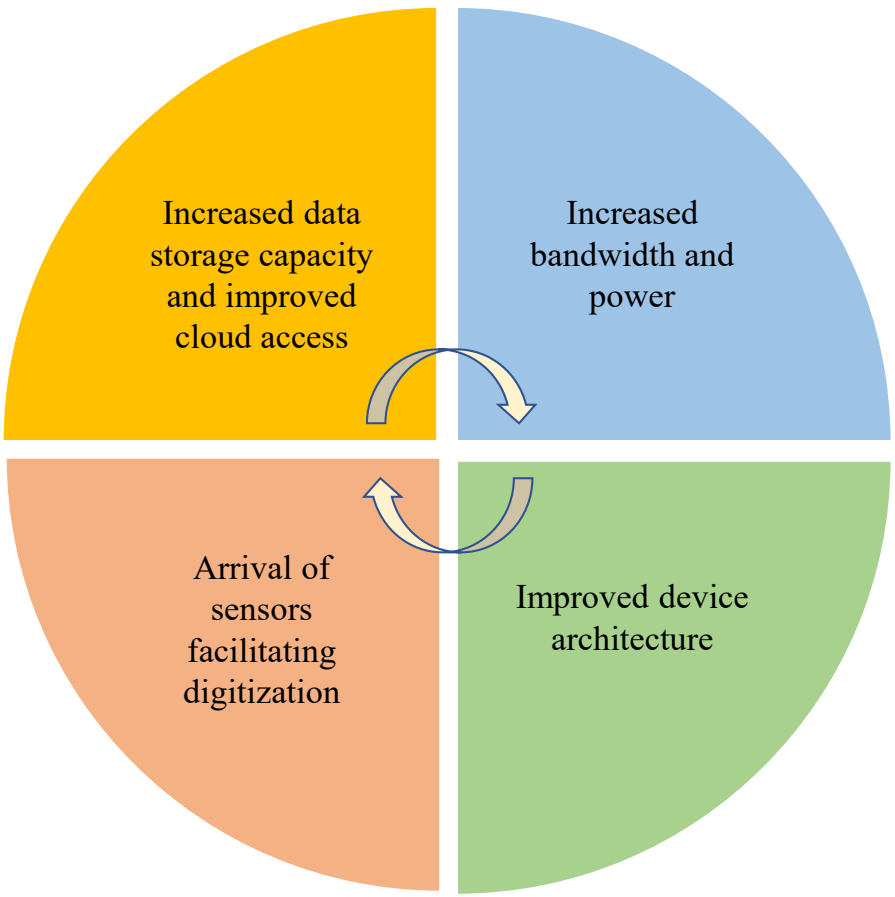
Several systematic approaches to assess the potential use of a biomarker in cervical cancer screening and management have been proposed (Arbyn et al., 2009; Wentzensen & Wacholder, 2013). Established guidelines for diagnostic research (the Standards for Reporting of Diagnostic Accuracy Studies [STARD] statement) have been adapted for technology development for cervical cancer screening (Arbyn et al., 2009). Five phases of technology evaluation are formally distinguished: (1) preclinical exploratory studies, (2) clinical validation studies, (3) retrospective biobank studies in the target population, (4) prospective screening studies, and (5) prospective intervention studies. Although this framework provides important guidance for technology development, not all of these steps are required for all technologies, and the sequence may vary depending on the clinical indication and the availability of suitable research studies. The evaluation of a technology must occur in the context of its potential use, because diagnostic accuracy requirements differ depending on whether the technology is used in screening, triage, or disease confirmation. Here, the term "emerging technology" is used when the discovery processes have been completed and the early steps of technology evaluation are under way (i.e. phases 1–3).

The process from discovery and development to clinical implementation is complex and involves many stakeholders, including researchers, industry, regulatory authorities, and professional societies that develop guidelines (Wentzensen & Silver, 2016). It can take a long time from initial discovery to clinical implementation. For example, HPV DNA testing was initially developed in the 1980s but did not enter clinical practice until 20 years later. The timeline from discovery to clinical practice is now shorter, because of the better understanding of the natural history of cervical cancer and the much accelerated technology development.

Because most discovered biomarkers do not make it into clinical practice, it is important to identify likely failures early in the evaluation process, enabling researchers to focus on the most promising leads (Wentzensen & Wacholder, 2013). The most important criterion for a biomarker is whether the test result will improve clinical management; if not, the test may be useless. Successful biomarker development usually relies on a commercial party to invest in assay development and regulatory approval. Therefore, barriers to bringing a promising biomarker into clinical practice may be the lack of intellectual property, or relatively limited clinical indication, which may result in too small a commercial market.

Of the molecular technologies summarized here, some were developed several years ago but have not been sufficiently validated for consideration of clinical use or have not been translated from the research setting to a commercially available test, for various reasons. Other novel technologies are rapidly progressing through the evaluation process, such as AI-based visual and cytological methods, as well as host and viral DNA methylation markers, which can be expected to appear in extensive clinical validation studies very soon.

Fig. 4.8 Pathway to the development of new technologies



Created by the Working Group.

4.6.1 Emerging technologies using artificial intelligence

AI is having an impact on many scientific disciplines, including medicine. As the power of computer software has increased, the size of the hardware has decreased, and as Internet bandwidth and electronic storage capacity have improved, it has become possible to deliver accurate image-recognition systems in very small, cloud-independent devices that incorporate comprehensive systems for management of clinical data and images (Fig. 4.8). Convolutional neural networks (CNNs) are commonly used for the analysis and classification of visual images; they are increasingly being used in medical diagnostics, such as in the classification of benign or malignant lung tumours (Hussein et al., 2017), in skin cancer (Esteva et al., 2017), in retinopathy (Ting et al., 2017), in the classification of colorectal polyps (Wei et al., 2020), in breast cancer (McKinney et al., 2020), and in the detection of cardiological abnormalities (Islam et al., 2017). Recently, these approaches have also been applied to automated and biomarker-enhanced cervical cytology (Schiffman et al., 2017; Wentzensen et al., 2021).

(a) AI-based automated visual evaluation

Even with adequate training and quality assurance measures in place, visual inspection of the cervix is a highly subjective procedure, including determining the adequacy of the examination, the type of the TZ, and the diagnostic impression. Furthermore, comprehensive training to the level of independent practice can take 6–18 months. A major but not exclusive part of this training is in image recognition, which to date has been learned largely within a live clinical setting. The concept of training a computer to recognize abnormality by "learning" the relevant features from a large image bank of known histopathology has obvious appeal. If that computational power can be harnessed in small, inexpensive, and user-friendly image-capture systems, the inadequacies of current visual examination methods could be addressed without the need for expensive training or adjunctive systems. As a laboratory-independent and reusable device, this technology could replace or complement current visual-based screening and triage approaches in LMICs. It may also negate the need for individual colposcopy expertise in screen-positive women who are not suitable for ablative treatment as part of a screen-and-treat protocol. AI can be used innovatively to train service providers and for quality control. Currently, no system has been properly evaluated in a live or real-world setting.

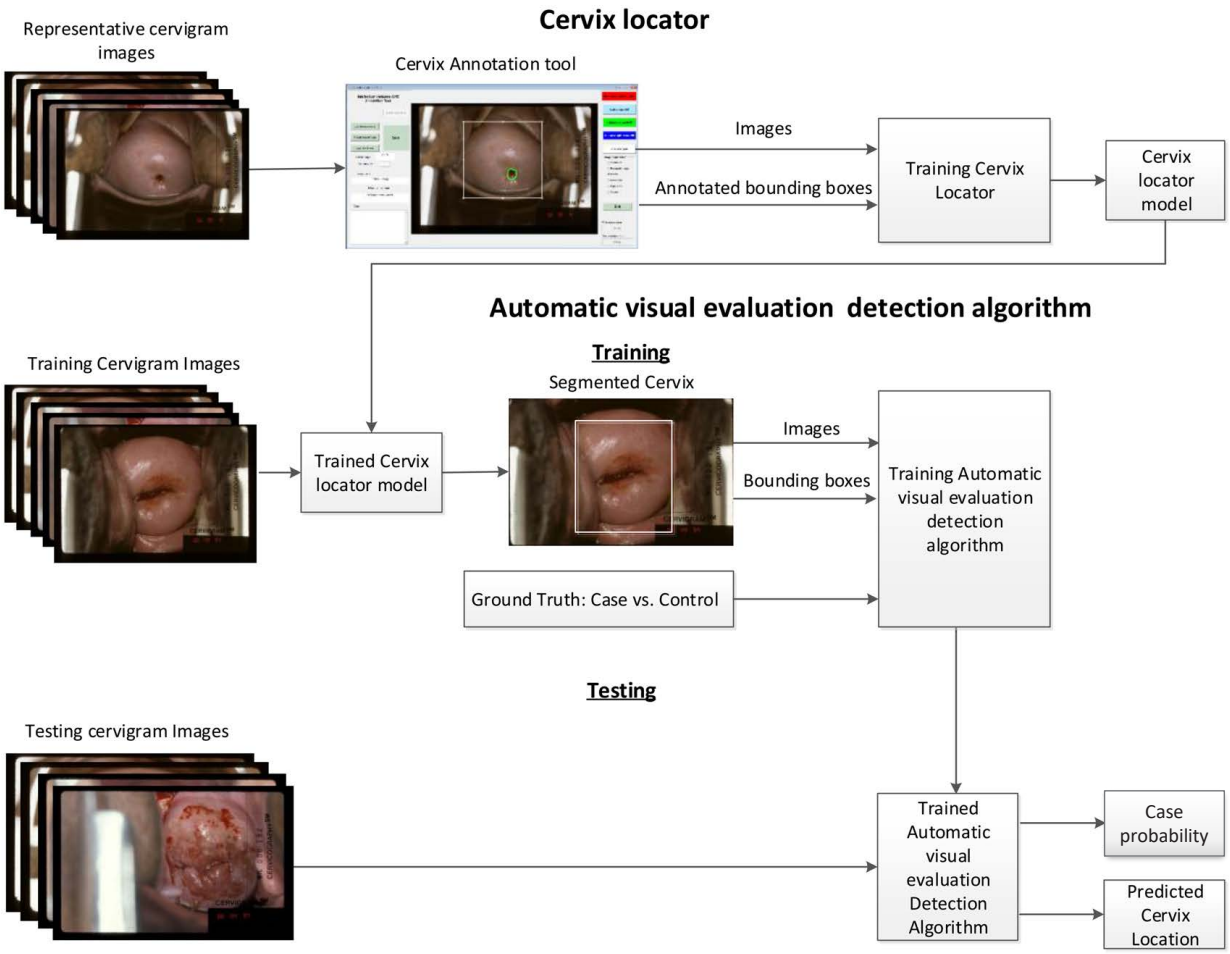
(i) Technical description

Training a model to discriminate between one image and another is now feasible, thanks to improved technology. Also, computing power has increased exponentially, and large, appropriately labelled image banks are available. Currently, for the detection of squamous cervical precancer, the clinically important discriminatory threshold is between normal or LSIL and HSIL. Therefore, algorithms in cervical precancer detection have focused on this dichotomous division. Training a CNN to discriminate between two distinct epithelial appearances within the squamous epithelium of the TZ involves exposing the model to a large series of adequate cervical images of known severity (i.e. supported by histopathology). Moreover, specific features on the cervical image may also be labelled by experts for a model to process. The CNN may then categorize cervical images into one of the two categories (≤ LSIL or HSIL) by outputting the probability that a given image belongs to either category.

During training, the CNN receives as inputs images from the training data set and adjusts its parameters to minimize the error between its predictions and the ground truth (i.e. colposcopically or histologically verified disease status) of the training set. Thus, the CNN is fitted to the training data set, learning the relevant features from the training data set, which enables it to increase the number of correct predictions. This process is illustrated in Fig. 4.9 (Hu et al., 2019). While the model is being trained on the training data set, the discriminative performance of the model is evaluated in a validation set. The purpose of the validation set is to evaluate the performance of the model on data that it has not been fitted to during the training process. Models with different selected hyperparameters can be trained in this way until a model that performs optimally on the validation set is determined. This yields a final trained model that can then be evaluated on a test set of images to assess its generalizability to predict cervical disease.

In general, the larger the training set, the higher the accuracy of the model. A viable model is often only as good as the quality of the images on which it is trained and the labels, or the robustness of the disease end-points, associated with these images. In many medical applications, there is often an imbalance between the number of images in each category; for example, in most cervical precancer image banks there are more images of ≤ LSIL than of HSIL. This imbalance can affect the training and validation process for the development of the model. The

Fig. 4.9 System architecture of the automated visual evaluation algorithm used by Hu et al. (2019)



From Hu et al. (2019).

scarcity of accurately labelled medical data, or robust disease end-points, with which to train CNNs for certain medical problems is a challenge to computational analysis. Although large image repositories may be available in some cases, relevant labelling of these images or information about the methods used to determine disease may be unclear or limited, leading to risk of disease misclassification. In addition, the quality of the available images depends on the sophistication of the image-capture system used. However, several specialized techniques (e.g. augmentation, transfer learning) can be used to address these issues and improve the performance of the model.

(ii) Performance of method

This technology may be appropriate for both screening and triage of screen-positive women. Early work using deep learning in cervical imagery has been encouraging (Xu et al., 2017). A deep-learning-based object detection method (Ren et al., 2017) was used to develop a visual evaluation algorithm for the detection of cervical precancer. Digitized cervigrams were collected as part of a population-based longitudinal cohort

study in 9406 women in Costa Rica; 241 of the women had histopathological confirmation of precancer (CIN2/3), and 38 had cancer over 7 years of follow-up in 1993–2001 (Hu et al., 2019). Despite limitations in image quality and images without full visualization of the squamocolumnar junction, the algorithm showed high accuracy for the identification of cervical precancers (Fig. 4.10). Automated visual evaluation of cervigrams collected at enrolment identified the cumulative number of cases of precancer or cancer with greater accuracy (AUC, 0.91; 95% CI, 0.89–0.93) than interpretation of the same images by a colposcopist (cervicography; AUC, 0.69; 95% CI, 0.63–0.74; $P < 0.0001$) or conventional cytology (AUC, 0.71; 95% CI = 0.65–0.77; $P < 0.0001$).

AI or deep-learning algorithms may be developed in different ways. Because the discriminative model "reads" images, the image-capture technique is relevant. Using this approach, Xue et al. (2020) developed an algorithm to interpret images captured by the smartphone-based MobileODT system. Automated visual evaluation can classify images of the cervix taken using smartphone camera image-capture systems. Alternatives to this approach include the development of a dedicated high-quality image-capture device that can capture multiple images to mimic a thorough colposcopic evaluation. Such systems can incorporate all the necessary computational power within a single device that is independent of the cloud; this makes them useful in low-resource settings. Both approaches have yet to be evaluated in the field.

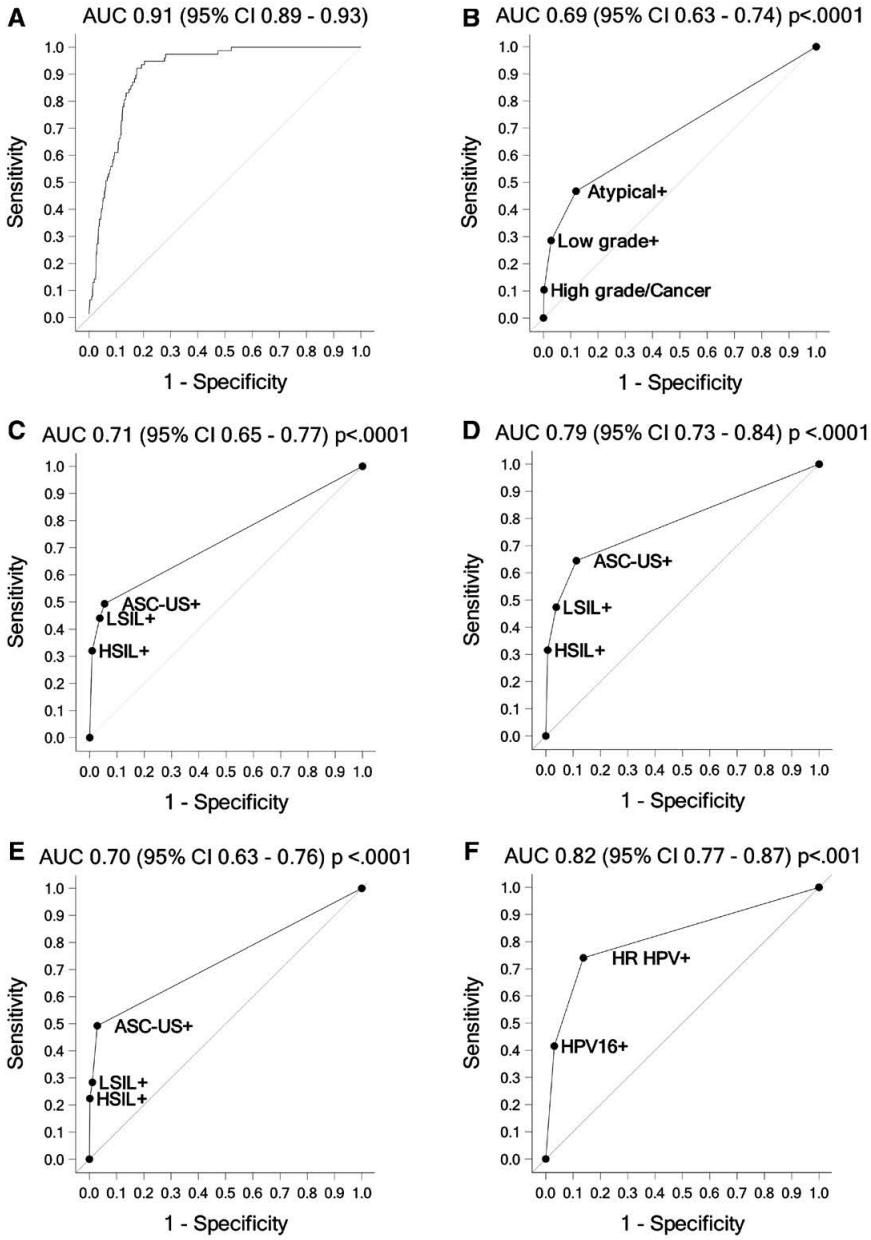
(b) Automated cytology technologies

Computer-assisted cytology systems have previously been developed for the reading of conventional or liquid-based cytology slides and are currently used in some settings. For the technical description and performance of these technologies, see Section 4.3.1(c). Recently, new AI-based approaches have been developed for automated evaluation of Pap cytology and dual-stain cytology.

A fully automated approach to evaluate Pap cytology was developed and validated in two studies in the USA. The training and validation data set included 1178 cervical cytology slides from HPV-positive women in Oklahoma who were referred for colposcopy for cytological abnormalities or for treatment of previously diagnosed precancer or cancer. The automated cytology algorithm achieved a performance for detection of CIN2+ (sensitivity, 0.91; specificity, 0.30) similar to that of conventional cytology with a threshold of ASC-US+ (sensitivity, 0.94; specificity, 0.30) (Schiffman et al., 2017). A subsequent study in 1839 HPV-positive women in the KPNC cohort, of whom 310 had precancer (181 with CIN2 and 129 with CIN3/AIS), similarly reported comparability of automated cytology and LBC with a threshold of ASC-US+ and LSIL+ (Yu et al., 2018).

Cytology with p16/Ki-67 dual staining (see Section 4.3.1(e)), which is used as a triage marker for HPV-positive women (see Section 4.4.7), can also be read by an automated system. A CNN deep-learning-based automated algorithm has been developed to evaluate p16/Ki-67 dual-stained slides (CYTOREADER software). The system uses a whole-slide scan followed by a machine-learning algorithm to detect and quantify p16/Ki-67 dual-stain-positive cells. A deep-learning classifier for automated dual-stained slides was compared with manual dual staining and conventional cytology for the detection of precancer in 602 women in Oklahoma who were referred for colposcopy, of whom 53 (8.8%) had CIN3+ (Wentzensen et al., 2021). The automated dual-staining algorithm had marginally lower positivity than manual dual staining (58% vs 63%; $P = 0.06$), with comparable sensitivity for the detection of CIN3+ (automated dual staining: 87%; 95% CI, 76–94%; manual dual staining: 87%; 95% CI, 76–94%; $P = 1.0$) and marginally higher specificity (automated

Fig. 4.10 ROC curve of automated visual evaluation of cervical images, and comparison of performance in identification of CIN2+



ROC-like curves are shown for the categorical variables for simple visual and statistical comparison with automated visual evaluation (two-sided χ^2 tests). The thresholds are listed on each curve, showing the sensitivity and 1 – specificity applicable to that threshold. Automated visual evaluation was as accurate as or more accurate than all of the screening tests used in the cohort study: (A) automated visual evaluation, (B) cervicography, (C) conventional cytology, (D) liquid-based cytology, (E) first-generation neural network-based cytology, and (F) MY09/MY11 PCR-based hrHPV testing.

ASC-US+. atypical squamous cells of undetermined significance or worse; AUC, area under the curve; CI, confidence interval; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; HPV, human papillomavirus; hrHPV, high-risk human papillomavirus; HSIL+, high-grade squamous intraepithelial lesion or worse; LSIL+, low-grade squamous intraepithelial lesion or worse; PCR, polymerase chain reaction; ROC, receiver operating characteristic.

From Hu et al. (2019).

dual staining: 46%; 95% CI, 41–51%; manual dual staining: 41%; 95% CI, 36–46%; $P = 0.07$). Similarly, in 3095 HPV-positive women undergoing routine cervical cancer screening in the KPNC cohort, of whom 218 (7.0%) had CIN3+, the test positivity of the automated dual-staining algorithm was significantly lower than that of manual dual staining or conventional cytology with a threshold of ASC-US+ (42%, 50%, and 60%, respectively), with comparable sensitivity (88%, 90%, and 86%, respectively) and higher specificity (62%, 53%, and 42%, respectively). The automated dual-staining algorithm led to a substantial reduction in the colposcopy referral rate compared with conventional cytology, paired with better disease detection, and provided additional risk stratification compared with manual dual staining in HPV-positive women.

4.6.2 Emerging molecular technologies

HPV-based testing may soon replace cytology as the primary screening method for cervical cancer in many parts of the world. However, the lower specificity of HPV DNA-based tests means that some screen-positive women are referred for colposcopy unnecessarily. Novel methods are required to identify which HPV-positive women need to be referred for colposcopy (Cuschieri et al., 2018). Although infection with carcinogenic HPV is necessary for the development of cervical cancer, other molecular changes occur with carcinogenic HPV infection, which result from DNA nucleotide mutations, structural genomic variations, or epigenetic alterations, such as DNA methylation (Steenbergen et al., 2014). Aberrant DNA methylation may help to distinguish non-progressive HPV infections from those that will progress to cervical cancer. It may thus be used as a strategy to triage HPV-positive women.

(a) DNA methylation

(i) Technical description

DNA methylation occurs after the addition of a methyl group to position 5 of the cytosine (C) ring immediately preceding a guanine (G) in the DNA sequence. It occurs mainly at CpG dinucleotide sites (C and G separated by one phosphate), known as CpG islands, which are present in about 60% of human genes (Laird, 2010). Controlled DNA methylation is essential for normal biological processes, such as the regulation of cellular processes including embryonic development, chromosomal instability, and protection from invading foreign viral DNA. However, aberrant DNA methylation can lead to alterations in the functions of gene products that regulate tumour suppression, DNA repair, apoptosis, metastasis, and invasion (Steenbergen et al., 2014; Lorincz, 2016). DNA methylation of some human genes and of the genome of hrHPV genotypes has been shown to be associated with increasing persistence of hrHPV genotypes (Mirabello et al., 2012), precancer (Wentzensen et al., 2009; Bierkens et al., 2013), and invasive cervical cancer (Bowden et al., 2019; Cook et al., 2019; Kelly et al., 2019). DNA methylation of more than 100 human genes and up to 12 carcinogenic HPV genotypes has been evaluated as a possible biomarker for the detection of cervical precancer and cancer using clinician-collected or self-collected cervical samples (Wentzensen et al., 2009; Lorincz, 2016).

(ii) Host DNA methylation

The most widely studied human gene DNA methylation targets have been evaluated as triage tests in HPV-positive women in cross-sectional, case–control, or convenience studies. Most studies evaluated the DNA methylation of the human genes *CADM1*, *MAL*, and *miR-124-2* in different combinations, and of *PAX-1*, *SOX-1*, *POU4F3*, and *FAM19A4*, alone or in combination with *miR-124-2*, for the detection of CIN2+

or CIN3+. Several studies evaluated the DNA methylation of the human gene *EPB41L3*, alone or in combination with DNA methylation of HPV16 (late coding regions L1 and L2), HPV18 (L2), HPV31 (L1), and HPV33 (L2), which is defined as the S5 classifier. The sensitivity and specificity of DNA methylation assays for the detection of prevalent CIN2+ have been shown to vary widely depending on the human gene target, the CpG targets of the gene studied, variations in the thresholds used to define methylation positivity, and the study design (Lorincz, 2016; Kelly et al., 2019).

RCTs comparing detection of CIN2+ in women undergoing testing with DNA methylation compared with cytology, and prospective studies evaluating baseline DNA methylation status to predict the risk of cervical cancer over time have been informative in clarifying the value of DNA methylation as a triage test.

In a non-inferiority RCT (Protection by Offering HPV Testing on Self-Sampled Cervicovaginal Specimens Trial 3 [PROHTECT-3]) in the Netherlands, HPV-positive women registered in the national cervical cancer screening programme who submitted a self-collected sample were randomly allocated to either triage with cytology (509 women) or triage with DNA methylation analysis of the *MAL* and *miR-124-2* genes (515 women) (Verhoef et al., 2014). Detection of CIN2+ with triage by methylation was non-inferior to that by cytology (17% vs 15%; RR, 1.19; 95% CI, 0.90–1.57), and the sensitivity for detection of CIN2+ was equivalent (adjusted sensitivity, 71%; 95% CI, 66–75% for both DNA methylation and cytology), although the sensitivity for detection of CIN3+ was slightly lower with DNA methylation (68%; 95% CI, 63–72%) than with cytology (75%; 95% CI, 70–79%). Also, because of a lower specificity to distinguish < CIN2, referral for colposcopy was more common in the methylation group than in the cytology group (55% vs 29%; $P < 0.0001$) (Verhoef et al., 2014). In a 14-year longitudinal study in 1040 HPV-positive women enrolled in the POBASCAM screening trial in the Netherlands, all of whom underwent testing with DNA methylation and cytology, a negative *FAM19A4/miR-124-2* methylation test indicated lower risk of cervical cancer incidence over a 14-year follow-up period compared with a negative cytology result (< ASC-US) at enrolment (risk ratio, 0.71; 95% CI, 0.16–1.40) (De Strooper et al., 2018).

Previous studies have shown high agreement between clinician-collected and self-collected samples and between lavage-based and brush-based self-collected samples for several human gene DNA methylation targets (Boers et al., 2014; De Strooper et al., 2016); this offers the possibility of conducting screening and triage on the same self-collected specimen.

(iii) Viral DNA methylation

DNA methylation of the early (E2) and late (L1 and L2) coding regions of the HPV viral genome has been reported to increase with increasing CIN grade for 12 carcinogenic HPV types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59 (Clarke et al., 2012; Wentzensen et al., 2012; Lorincz et al., 2013; Mirabello et al., 2013; Bowden et al., 2019). The diagnostic accuracy of DNA methylation of HPV genotypes, alone or in various combinations, has been evaluated for detection of CIN2+. In a meta-analysis of seven studies evaluating DNA methylation of the E2, L1, and/or L2 coding regions of HPV16 in HPV16-positive women, the pooled sensitivity for detection of CIN2+ was 74% (95% CI, 57–85%) and the pooled specificity was 73% (95% CI, 66–79%), although there was significant heterogeneity in the observed estimates, because of differences in the CpG sites targeted (Kelly et al., 2019). A second, independent meta-analysis on the diagnostic accuracy of the HPV16 *L1* and/or *L2* genes in 10 studies reported similar findings, with a pooled sensitivity of 77% (95% CI, 63–87%) and a pooled specificity of 64% (95% CI, 55–71%) (Bowden et al., 2019).

The addition of HPV type-specific methylation (HPV types 16, 18, 31, and 33) to a human gene target (*EPB41L3*) as part of the S5 classifier enables testing in all women, irrespective of HPV type positivity. In three studies conducted in HPV-positive women in Canada, Colombia, and the United Kingdom, the sensitivity of the S5 classifier varied from 74% to 82% for detection of CIN2+ and from 84% to 93% for detection of CIN3+, suggesting that the combination of viral and host gene targets may increase detection of CIN2+/CIN3+ (Lorincz et al., 2016; Cook et al., 2019; Ramírez et al., 2021). However, the specificity for < CIN2 varied from 35% to 65%. Compared with either cytology with a threshold of ASC-US+ or HPV16/18 partial genotyping, the S5 classifier had a consistently higher sensitivity for the detection of CIN2+ or CIN3+ but a lower specificity (Lorincz et al., 2016; Cook et al., 2019; Ramírez et al., 2021).

A multiplex DNA methylation test targeting the L1/L2 regions of a wider range of HPV types (HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59) was evaluated in a case–control study in 299 women with precancer (CIN3/AIS) and 360 women who had normal cytology but who were positive for any one of the targeted HPV types (i.e. 30 controls for each of the 12 carcinogenic HPV types evaluated) (Clarke et al., 2018). Methylation was positively associated with CIN3/AIS for all 12 types. The diagnostic accuracy of the 12-type DNA methylation assay was simulated by applying type-specific sensitivity and specificity estimates for the DNA methylation test to a population of 30 000 women using data from a cohort of women undergoing routine cervical screening in the USA. The simulated sensitivity and specificity of the 12-type DNA methylation assay were 80% and 66%, respectively; both were higher than for cytology with a threshold of ASC-US+ (77% and 54%, respectively).

(b) Detection of HPV E6 oncoprotein

Elevated expression of the HPV oncoproteins E6 and E7 is associated with the development of HPV-associated cervical cancer. E6 oncoprotein from HPV16/18/45 can be detected by the Onco*E6* test (Wentzensen et al., 2016). Zhao et al. (2013) reported the test performance when E6 oncoprotein was used as a primary screening method. Another study in China assessed the test performance of E6 oncoprotein for the detection of CIN3+ as triage for HPV-positive women (Qiao et al., 2014). The sensitivity of E6 oncoprotein from HPV16/18/45 was about 50% and the specificity was more than 90% in both clinician-collected and self-collected samples. Compared with HPV16/18/45 DNA testing, the sensitivity was lower but the specificity was higher.

A recent study reported the cumulative incidence of CIN3+ in 1742 women at 10-year follow-up (Dong et al., 2020). The cumulative incidence of CIN3+ was higher in women harbouring methylation at six sites (CpG 5602, 6650, 7034, 7461, 31, and 37) with and without E6 oncoprotein than in women with abnormal cytology. For triage of HPV16-positive women with detection of CIN3+, the sensitivity of E6 oncoprotein was lower than that of cytology (57.1% vs 92.9%), but the specificity was higher (86.5% vs 43.2%). A higher AUC was obtained with the methylation test at the six sites (0.82; 95% CI, 0.69–0.91) than with E6 oncoprotein detection (0.72; 95% CI, 0.58–0.82) and with cytology (0.68; 95% CI, 0.54–0.80).

References

Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE (2009). How to evaluate emerging technologies in cervical cancer screening? *Int J Cancer*. 125(11):2489–96. doi:10.1002/ijc.24774 PMID:19626591

Bierkens M, Hesselink AT, Meijer CJLM, Heideman DAM, Wisman GBA, van der Zee AGJ, et al. (2013). *CADM1* and *MAL* promoter methylation levels in

hrHPV-positive cervical scrapes increase proportional to degree and duration of underlying cervical disease. *Int J Cancer*. 133(6):1293–9. doi:10.1002/ijc.28138 PMID:23456988

Boers A, Bosgraaf RP, van Leeuwen RW, Schuuring E, Heideman DAM, Massuger LFAG, et al. (2014). DNA methylation analysis in self-sampled brush material as a triage test in hrHPV-positive women. *Br J Cancer*. 111(6):1095–101. doi:10.1038/bjc.2014.392 PMID:25032730

Bowden SJ, Kalliala I, Veroniki AA, Arbyn M, Mitra A, Lathouras K, et al. (2019). The use of human papillomavirus DNA methylation in cervical intraepithelial neoplasia: a systematic review and meta-analysis. *EBioMedicine*. 50:246–59. doi:10.1016/j.ebiom.2019.10.053 PMID:31732479

Clarke MA, Gradissimo A, Schiffman M, Lam J, Sollecito CC, Fetterman B, et al. (2018). Human papillomavirus DNA methylation as a biomarker for cervical precancer: consistency across 12 genotypes and potential impact on management of HPV-positive women. *Clin Cancer Res*. 24(9):2194–202. doi:10.1158/1078-0432.CCR-17-3251 PMID:29420222

Clarke MA, Wentzensen N, Mirabello L, Ghosh A, Wacholder S, Harari A, et al. (2012). Human papillomavirus DNA methylation as a potential biomarker for cervical cancer. *Cancer Epidemiol Biomarkers Prev*. 21(12):2125–37. doi:10.1158/1055-9965.EPI-12-0905 PMID:23035178

Cook DA, Krajden M, Brentnall AR, Gondara L, Chan T, Law JH, et al. (2019). Evaluation of a validated methylation triage signature for human papillomavirus positive women in the HPV FOCAL cervical cancer screening trial. *Int J Cancer*. 144(10):2587–95. doi:10.1002/ijc.31976 PMID:30412281

Cuschieri K, Ronco G, Lorincz A, Smith L, Ogilvie G, Mirabello L, et al. (2018). Eurogin roadmap 2017: triage strategies for the management of HPV-positive women in cervical screening programs. *Int J Cancer*. 143(4):735–45. doi:10.1002/ijc.31261 PMID:29341110

De Strooper LMA, Berkhof J, Steenbergen RDM, Lissenberg-Witte BI, Snijders PJF, Meijer CJLM, et al. (2018). Cervical cancer risk in HPV-positive women after a negative *FAM19A4/mir124-2* methylation test: a post hoc analysis in the POBASCAM trial with 14 year follow-up. *Int J Cancer*. 143(6):1541–8. doi:10.1002/ijc.31539 PMID:29663363

De Strooper LMA, Verhoef VMJ, Berkhof J, Hesselink AT, de Bruin HME, van Kemenade FJ, et al. (2016). Validation of the *FAM19A4/mir124-2* DNA methylation test for both lavage- and brush-based self-samples to detect cervical (pre)cancer in HPV-positive women. *Gynecol Oncol*. 141(2):341–7. doi:10.1016/j.ygyno.2016.02.012 PMID:26921784

Dong L, Zhang L, Hu SY, Feng RM, Zhao XL, Zhang Q, et al. (2020). Risk stratification of HPV 16 DNA methylation combined with E6 oncoprotein in cervical cancer screening: a 10-year prospective cohort study. *Clin Epigenetics*. 12(1):62. doi:10.1186/s13148-020-00853-1 PMID:32381054

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 542(7639):115–8. doi:10.1038/nature21056 PMID:28117445

Hu L, Bell D, Antani S, Xue Z, Yu K, Horning MP, et al. (2019). An observational study of deep learning and automated evaluation of cervical images for cancer screening. *J Natl Cancer Inst*. 111(9):923–32. doi:10.1093/jnci/djy225 PMID:30629194

Hussein S, Cao K, Song Q, Bagci U (2017). Risk stratification of lung nodules using 3D CNN-based multi-task learning. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap PT, et al., editors. Information processing in medical imaging. IPMI 2017. (Lecture Notes in Computer Science, Volume 10265). Cham, Switzerland: Springer; pp. 249–60. doi:10.1007/978-3-319-59050-9_20

Islam MT, Aowal MA, Tahseen Minhaz A, Ashraf K (2017). Abnormality detection and localization in chest X-rays using deep convolutional neural networks. Ithaca (NY), USA: Cornell University. Available from: https://arxiv.org/pdf/1705.09850.pdf.

Kelly H, Benavente Y, Pavon MA, De Sanjose S, Mayaud P, Lorincz AT (2019). Performance of DNA methylation assays for detection of high-grade cervical intraepithelial neoplasia (CIN2+): a systematic review and meta-analysis. *Br J Cancer*. 121(11):954–65. doi:10.1038/s41416-019-0593-4 PMID:31616037

Laird PW (2010). Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet*. 11(3):191–203. doi:10.1038/nrg2732 PMID:20125086

Lorincz AT (2016). Virtues and weaknesses of DNA methylation as a test for cervical cancer prevention. *Acta Cytol*. 60(6):501–12. doi:10.1159/000450595 PMID:27806357

Lorincz AT, Brentnall AR, Scibior-Bentkowska D, Reuter C, Banwait R, Cadman L, et al. (2016). Validation of a DNA methylation HPV triage classifier in a screening sample. *Int J Cancer*. 138(11):2745–51. doi:10.1002/ijc.30008 PMID:26790008

Lorincz AT, Brentnall AR, Vasiljević N, Scibior-Bentkowska D, Castanon A, Fiander A, et al. (2013). HPV16 L1 and L2 DNA methylation predicts high-grade cervical intraepithelial neoplasia in women with mildly abnormal cervical cytology. *Int J Cancer*. 133(3):637–44. doi:10.1002/ijc.28050 PMID:23335178

McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. (2020). International evaluation of an AI system for breast cancer screening.

Nature. 577(7788):89–94. doi:10.1038/s41586-019-1799-6 PMID:31894144

Mirabello L, Schiffman M, Ghosh A, Rodriguez AC, Vasiljevic N, Wentzensen N, et al. (2013). Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia. *Int J Cancer*. 132(6):1412–22. doi:10.1002/ijc.27750 PMID:22847263

Mirabello L, Sun C, Ghosh A, Rodriguez AC, Schiffman M, Wentzensen N, et al. (2012). Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population. *J Natl Cancer Inst*. 104(7):556–65. doi:10.1093/jnci/djs135 PMID:22448030

Qiao YL, Jeronimo J, Zhao FH, Schweizer J, Chen W, Valdez M, et al. (2014). Lower cost strategies for triage of human papillomavirus DNA-positive women. *Int J Cancer*. 134(12):2891–901. doi:10.1002/ijc.28616 PMID:24248915

Ramírez AT, Sánchez GI, Nedjai B, Agudelo MC, Brentnall AR, Cuschieri K, et al.; ASC-US-COL Trial Group (2021). Effective methylation triage of HPV positive women with abnormal cytology in a middle-income country. *Int J Cancer*. 148(6):1383–93. doi:10.1002/ijc.33314 PMID:33006394

Ren S, He K, Girshick R, Sun J (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell*. 39(6):1137–49. doi:10.1109/TPAMI.2016.2577031 PMID:27295650

Schiffman M, Yu K, Zuna R, Terence Dunn S, Zhang H, Walker J, et al. (2017). Proof-of-principle study of a novel cervical screening and triage strategy: computer-analyzed cytology to decide which HPV-positive women are likely to have ≥CIN2. *Int J Cancer*. 140(3):718–25. doi:10.1002/ijc.30456 PMID:27696414

Steenbergen RDM, Snijders PJF, Heideman DAM, Meijer CJLM (2014). Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat Rev Cancer*. 14(6):395–405. doi:10.1038/nrc3728 PMID:24854082

Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A, et al. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 318(22):2211–23. doi:10.1001/jama.2017.18152 PMID:29234807

Verhoef VMJ, Bosgraaf RP, van Kemenade FJ, Rozendaal L, Heideman DAM, Hesselink AT, et al. (2014). Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): a randomised controlled non-inferiority trial. *Lancet Oncol*. 15(3):315–22. doi:10.1016/S1470-2045(14)70019-1 PMID:24529697

Wei JW, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, et al. (2020). Evaluation of a deep neural network for automated classification of colorectal polyps on histopathologic slides. *JAMA Netw Open*. 3(4):e203398. doi:10.1001/jamanetworkopen.2020.3398 PMID:32324237

Wentzensen N, Lahrmann B, Clarke MA, Kinney W, Tokugawa D, Poitras N, et al. (2021). Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical cancer screening. *J Natl Cancer Inst*. 113(1):72–9. doi:10.1093/jnci/djaa066 PMID:32584382

Wentzensen N, Schiffman M, Palmer T, Arbyn M (2016). Triage of HPV positive women in cervical cancer screening. *J Clin Virol*. 76(Suppl 1):S49–55. doi:10.1016/j.jcv.2015.11.015 PMID:26643050

Wentzensen N, Sherman ME, Schiffman M, Wang SS (2009). Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecol Oncol*. 112(2):293–9. doi:10.1016/j.ygyno.2008.10.012 PMID:19054549

Wentzensen N, Silver MI (2016). Biomarkers for cervical cancer prevention programs: the long and winding road from discovery to clinical use. *J Low Genit Tract Dis*. 20(3):191–4. doi:10.1097/LGT.0000000000000231 PMID:27243141

Wentzensen N, Sun C, Ghosh A, Kinney W, Mirabello L, Wacholder S, et al. (2012). Methylation of HPV18, HPV31, and HPV45 genomes and cervical intraepithelial neoplasia grade 3. *J Natl Cancer Inst*. 104(22):1738–49. doi:10.1093/jnci/djs425 PMID:23093560

Wentzensen N, Wacholder S (2013). From differences in means between cases and controls to risk stratification: a business plan for biomarker development. *Cancer Discov*. 3(2):148–57. doi:10.1158/2159-8290.CD-12-0196 PMID:23299199

Xu T, Zhang H, Xin C, Kim E, Long LR, Xue Z, et al. (2017). Multi-feature based benchmark for cervical dysplasia classification evaluation. *Pattern Recognit*. 63:468–75. doi:10.1016/j.patcog.2016.09.027 PMID:28603299

Xue Z, Novetsky AP, Einstein MH, Marcus JZ, Befano B, Guo P, et al. (2020). A demonstration of automated visual evaluation of cervical images taken with a smartphone camera. *Int J Cancer*. 147(9):2416–23. doi:10.1002/ijc.33029 PMID:32356305

Yu K, Hyun N, Fetterman B, Lorey T, Raine-Bennett TR, Zhang H, et al. (2018). Automated cervical screening and triage, based on HPV testing and computer-interpreted cytology. *J Natl Cancer Inst*. 110(11):1222–8. doi:10.1093/jnci/djy044 PMID:29659930

Zhao FH, Jeronimo J, Qiao YL, Schweizer J, Chen W, Valdez M, et al. (2013). An evaluation of novel, lower-cost molecular screening tests for human papillomavirus in rural China. *Cancer Prev Res (Phila)*. 6(9):938–48. doi:10.1158/1940-6207.CAPR-13-0091 PMID:23878179