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
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
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 4.6 Systematic reviews of studies of test performance of manual diagnosis compared with automated screening

<table>
<thead>
<tr>
<th>Test threshold</th>
<th>Manual</th>
<th>Automated (ThinPrep)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%) (95% CI)</td>
<td>Specificity (%) (95% CI)</td>
<td>Sensitivity (%) (95% CI)</td>
</tr>
<tr>
<td>ASC-US</td>
<td>75.6 (72.2–78.8)</td>
<td>97.6 (97.2–97.9)</td>
<td>82.0 (78.8–84.8)</td>
</tr>
<tr>
<td>LSIL</td>
<td>79.7 (75.3–83.7)</td>
<td>99.0 (98.8–99.2)</td>
<td>79.2 (74.7–83.2)</td>
</tr>
<tr>
<td>HSIL</td>
<td>74.1 (66.0–81.2)</td>
<td>99.4 (99.2–99.6)</td>
<td>79.9 (72.2–86.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test threshold</th>
<th>Manual</th>
<th>Automated (FocalPoint)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
<td>Sensitivity (%)</td>
</tr>
<tr>
<td>ASC-US</td>
<td>82.6</td>
<td>82.7</td>
<td>81.1</td>
</tr>
<tr>
<td>LSIL</td>
<td>76.4</td>
<td>90.6</td>
<td>86.1</td>
</tr>
<tr>
<td>HSIL</td>
<td>65.7</td>
<td>97.7</td>
<td>85.3</td>
</tr>
</tbody>
</table>

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

* Reference standards in both studies were defined as the diagnosis of cytology carried out by experts in each study.
### 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)
  - Specimen rejected/not processed (specify reason)
  - 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
  - Squamous metaplasia
  - Keratotic changes
  - Tubal metaplasia
  - Atrophy
  - Pregnancy-associated changes
- Reactive cellular changes associated with:
  - Inflammation (includes typical repair)
    - Lymphocytic (follicular) cervicitis
  - Radiation
  - 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
  - Of undetermined significance
  - Cannot exclude HSIL
- LSIL (encompassing: HPV/mild dysplasia/CIN1)
  - HSIL (encompassing: moderate and severe dysplasia, CIS; CIN2 and CIN3)
    - With features suspicious for invasion (if invasion is suspected)
  - Squamous cell carcinoma
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\(^{\text{INK4a}}\) (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; Benevelo 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

<table>
<thead>
<tr>
<th>Reference Country</th>
<th>Outcome: CIN2+</th>
<th></th>
<th></th>
<th>Outcome: CIN3+</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%) (95% CI)</td>
<td>Specificity (%) (95% CI)</td>
<td>Sensitivity (%) (95% CI)</td>
<td>Specificity (%) (95% CI)</td>
<td>Sensitivity (%) (95% CI)</td>
<td>Specificity (%) (95% CI)</td>
</tr>
<tr>
<td>p16/Ki-67 dual staining</td>
<td>Cytology (ASC-US+)</td>
<td>HPV16/18 genotyping</td>
<td>p16/Ki-67 dual staining</td>
<td>Cytology (ASC-US+)</td>
<td>HPV16/18 genotyping</td>
<td>p16/Ki-67 dual staining</td>
</tr>
<tr>
<td>Petry et al. (2011) Germany</td>
<td>91.9 (78.1–98.3)</td>
<td>NA</td>
<td>NA</td>
<td>82.1 (72.9–89.2)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Wentzensen et al. (2012) USA</td>
<td>85.5 (77.8–90.9)</td>
<td>NA</td>
<td>47.6 (38.6–56.7)</td>
<td>59.4 (53.3–65.1)</td>
<td>NA</td>
<td>80.8 (75.5–85.2)</td>
</tr>
<tr>
<td>Wentzensen et al. (2015) USA</td>
<td>83.4 (77.1–88.6)</td>
<td>76.6 (69.6–82.6)</td>
<td>NA</td>
<td>58.9 (56.2–61.6)</td>
<td>49.6 (46.9–52.3)</td>
<td>NA</td>
</tr>
<tr>
<td>Gustinucci et al. (2016) Italy</td>
<td>87.6 (75.7–93.6)</td>
<td>77.6 (65.3–86.7)</td>
<td>47.0 (34.0–58.9)</td>
<td>74.9 (69.0–79.0)</td>
<td>72.5 (67.2–77.2)</td>
<td>77.9 (72.8–82.0)</td>
</tr>
<tr>
<td>Wright et al. (2017) USA</td>
<td>70.3 (65.3–74.9)</td>
<td>51.8 (46.5–58.3)</td>
<td>NA</td>
<td>75.6 (74.0–77.1)</td>
<td>76.1 (74.6–77.7)</td>
<td>NA</td>
</tr>
<tr>
<td>Stanczuk et al. (2017) United Kingdom</td>
<td>85.0 (73.4–92.9)</td>
<td>68.3 (55.0–79.7)</td>
<td>61.7 (48.2–73.9)</td>
<td>76.7 (71.1–81.8)</td>
<td>89.1 (84.7–92.7)</td>
<td>70.5 (64.6–76.0)</td>
</tr>
<tr>
<td>Wentzensen et al. (2019) USA</td>
<td>88.6 (84.5–92.6)</td>
<td>84.3 (79.7–89.0)</td>
<td>NA</td>
<td>53.1 (51.3–54.9)</td>
<td>42.9 (41.1–44.6)</td>
<td>NA</td>
</tr>
<tr>
<td>Stoler et al. (2020) USA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>86.0</td>
</tr>
<tr>
<td>Hu et al. (2020) China</td>
<td>63.5 (54.4–71.9)</td>
<td>61.9 (52.8–70.4)</td>
<td>61.9 (52.8–70.4)</td>
<td>85.3 (82.5–87.8)</td>
<td>80.0 (76.9–82.9)</td>
<td>72.4 (68.9–75.6)</td>
</tr>
<tr>
<td>Jiang et al. (2020) China</td>
<td>75.0 (50.9–91.3)</td>
<td>NA</td>
<td>NA</td>
<td>50.3 (41.9–58.8)</td>
<td>NA</td>
<td>83.3 (35.9–99.6)</td>
</tr>
</tbody>
</table>

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.
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).

Table 4.9 Comparison of guidelines for quality assurance for cytology

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sample collection</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Organization (staff, workload)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Material requirement</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quality management</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Terminology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Management of abnormal cytology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Follow-up</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Laboratory performance indicators</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quality improvement (audit)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

† Arbyn et al. (2008a).
§ National Pathology Accreditation Advisory Council (2019).
Table compiled by the Working Group.
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

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

CI, confidence interval; HR, hazard ratio.
<sup>a</sup> Rates and hazard ratios have been adjusted for age.
<sup>b</sup> 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

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

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

<table>
<thead>
<tr>
<th>Reference Country</th>
<th>Cohort description: no. of women, screening period, source of screening data, and source of follow-up data</th>
<th>Established programme: year of start, screening age, screening interval</th>
<th>Accrual and follow-up periods Person-years</th>
<th>Cervical cancer or precancer end-point, and incidence or mortality age ranges</th>
<th>No. of cases or deaths</th>
<th>Cervical cancer incidence or mortality RR (95% CI)*</th>
<th>Adjustments</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odongua et al. (2007) Republic of Korea</td>
<td>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</td>
<td>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</td>
<td>1992–2004 12 yr Person-yr not given</td>
<td>Cervical cancer incidence and mortality Age range at enrolment, 30–95 yr</td>
<td>2523 cases 209 deaths</td>
<td>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)]</td>
<td>Age, BMI, smoking status, alcohol consumption, menarche, parity</td>
<td>Only compared unscreened women with screened women with normal results, not all screening. Not clear that diagnostic smears were excluded</td>
</tr>
<tr>
<td>Reference Country</td>
<td>Cohort description: no. of women, screening period, source of screening data, and source of follow-up data</td>
<td>Established programme: year of start, screening age, screening interval</td>
<td>Accrual and follow-up periods Person-years</td>
<td>Cervical cancer or precancer end-point, and incidence or mortality age ranges</td>
<td>No. of cases or deaths</td>
<td>Cervical cancer incidence or mortality RR (95% CI)*</td>
<td>Adjustments</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Jun et al. (2009) Republic of Korea</td>
<td>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</td>
<td>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</td>
<td>1995–2002 Average follow-up time: 6.5 yr 1,657,130.4 person-yr</td>
<td>Incidence of invasive cervical cancer and CIS of the cervix Age ≥ 20 yr</td>
<td>248 cases of invasive cervical cancer 346 cases of CIS of the cervix</td>
<td>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)</td>
<td>Age, smoking status, alcohol consumption</td>
<td>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]</td>
</tr>
<tr>
<td>Reference Country</td>
<td>Cohort description: no. of women, screening period, source of screening data, and source of follow-up data</td>
<td>Established programme: year of start, screening age, screening interval</td>
<td>Accrual and follow-up periods Person-years</td>
<td>Cervical cancer or precancer end-point, and incidence or mortality age ranges</td>
<td>No. of cases or deaths</td>
<td>Cervical cancer incidence or mortality RR (95% CI)*</td>
<td>Adjustments</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>--------------------------------------------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Dugué et al. (2014) Denmark</td>
<td>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.</td>
<td>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</td>
<td>1998–2010 Person-yr not given</td>
<td>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)</td>
<td>No. of cervical cancer deaths: Never screened, 274 Irregularly screened, 152 Regularly screened, 237</td>
<td>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)]</td>
<td>Adjusted for age by using attained age as time scale in Cox proportional hazards regression</td>
<td>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]</td>
</tr>
<tr>
<td>Reference</td>
<td>Cohort description: no. of women, screening period, source of screening data, and source of follow-up data</td>
<td>Established programme: year of start, screening age, screening interval</td>
<td>Accrual and follow-up periods Person-years</td>
<td>Cervical cancer or precancer end-point, and incidence or mortality age ranges</td>
<td>No. of cases or deaths</td>
<td>Cervical cancer incidence or mortality RR (95% CI)*</td>
<td>Adjustments</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Wang et al. (2017) Sweden</td>
<td>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, or 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.</td>
<td>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.</td>
<td>Median follow-up time: unscreened, 10.6 yr; screened, 11.4 yr; overall, 10.9 yr Person-yr not given</td>
<td>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</td>
<td>868 cases of cervical cancer diagnosed at age 61–80 yr</td>
<td>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)</td>
<td>Education level, birth cohort Sensitivity analysis included parity and lifetime diagnosis of COPD as a proxy for smoking status</td>
<td>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.</td>
</tr>
<tr>
<td>Reference Country</td>
<td>Cohort description: no. of women, screening period, source of screening data, and source of follow-up data</td>
<td>Established programme: year of start, screening age, screening interval</td>
<td>Accrual and follow-up periods Person-years</td>
<td>Cervical cancer or precancer end-point, and incidence or mortality age ranges</td>
<td>No. of cases or deaths</td>
<td>Cervical cancer incidence or mortality RR (95% CI)</td>
<td>Adjustments</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Pankakoski et al. (2019) Finland</td>
<td>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 unininvited at 65 yr were not Helsinki residents, so had different underlying risk)</td>
<td>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</td>
<td>1991–2014 Median, 11.1 person-yr</td>
<td>Incidence-based mortality risk ratio of cervical cancer for women invited to routine screening at age 65 yr compared with those not invited</td>
<td>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</td>
<td>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)</td>
<td>Area of residence (background risk of cervical cancer)</td>
<td>Helsinki area was using cytology. Some areas were using HPV testing Unable to adjust for individual-level hysterectomy</td>
</tr>
</tbody>
</table>
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; Niimenen 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, 0.65; 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, 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 enrollees 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,
Cervical cancer screening

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

<table>
<thead>
<tr>
<th>Reference Trial, country</th>
<th>Randomization</th>
<th>No. of women</th>
<th>Population</th>
<th>Age at entry (years)</th>
<th>LBC procedure</th>
<th>Reference standard</th>
<th>Blinding of histological assessment?</th>
<th>Reported end-points</th>
<th>Long-term outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obwegeser &amp; Brack (2001) Switzerland</td>
<td>Individual Conv.: 1002 Conv.: 997</td>
<td>Opportunistic screening</td>
<td>15–≥ 70</td>
<td>ThinPrep 2000</td>
<td>Colposcopy for women with HSIL cytology; for ASC-US and LSIL, follow-up was mostly incomplete</td>
<td>No</td>
<td>CIN2+</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Taylor et al. (2006) South Africa</td>
<td>No; practice rotating every 6 mo</td>
<td>High-risk population</td>
<td>35–65</td>
<td>ThinPrep 2000</td>
<td>Colposcopy for all women</td>
<td>Yes</td>
<td>CIN2+ CIN3+</td>
<td>Not possible. Women were all referred for colposcopy</td>
<td></td>
</tr>
<tr>
<td>Ronco et al. (2007) NTCC, Italy</td>
<td>Individual Conv.: 22 466 Conv.: 22 708</td>
<td>Screening</td>
<td>25–60</td>
<td>ThinPrep</td>
<td>Colposcopy for all positive</td>
<td>CIN reviewed blindly</td>
<td>CIN2+ CIN3+</td>
<td>Not possible. Women were managed according to HPV test results</td>
<td></td>
</tr>
<tr>
<td>Strander et al. (2007) Sweden</td>
<td>Randomized per week of appointment Conv.: 8810 Conv.: 4674</td>
<td>Screening</td>
<td>23–60</td>
<td>ThinPrep 2000</td>
<td>Referral as routine practice; histology searched through registries</td>
<td>Yes</td>
<td>CIN2+</td>
<td>Cumulative incidence up to 3 yr and 7 mo</td>
<td></td>
</tr>
<tr>
<td>Sykes et al. (2008) New Zealand</td>
<td>Individual Conv.: 453 Conv.: 451</td>
<td>Women in colposcopy clinics</td>
<td>16–75</td>
<td>SurePath</td>
<td>Colposcopy-guided biopsy</td>
<td>No</td>
<td>CIN2+</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Maccallini et al. (2008) Italy</td>
<td>Individual Conv.: 4299 Conv.: 4355</td>
<td>Screening</td>
<td>25–64</td>
<td>ThinPrep</td>
<td>Colposcopy for all positive</td>
<td>No</td>
<td>CIN2+ CIN3+</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Siebers et al. (2008, 2009) NETHCON, Netherlands</td>
<td>Cluster RCT; family practice as randomization unit Conv.: 40 562 Conv.: 49 222</td>
<td>Screening</td>
<td>25–60</td>
<td>ThinPrep 3000</td>
<td>Referral as routine practice. All follow-up tests blindly reviewed</td>
<td>Yes</td>
<td>CIN2+ CIN3+</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Klug et al. (2013) Germany</td>
<td>Randomized per week of visit Conv.: 9 352 Conv.: 11 555</td>
<td>Opportunistic screening</td>
<td>≥ 20</td>
<td>ThinPrep with/without Imaging System</td>
<td>Colposcopy for all women with LSIL+</td>
<td>No</td>
<td>CIN2+ CIN3+</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Reference Country</th>
<th>Age (years)</th>
<th>Threshold</th>
<th>Total number</th>
<th>Detection rate (%)</th>
<th>PPV (%)</th>
<th>Unsatisfactory cytology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conv.</td>
<td>LBC</td>
<td>RR (95% CI)</td>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>Obwegeser &amp; Brack (2001)</td>
<td>15–70</td>
<td>ASC-US</td>
<td>1002</td>
<td>997</td>
<td>0.92 (0.41–2.07)</td>
<td>NA</td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor et al. (2006) South Africa</td>
<td>35–65</td>
<td>ASC-US</td>
<td>2444</td>
<td>3114</td>
<td>0.81 (0.54–1.21)</td>
<td>0.67 (0.39–1.14)</td>
</tr>
<tr>
<td>Strander et al. (2007) Sweden</td>
<td>23–60</td>
<td>LSIL</td>
<td>8810</td>
<td>4674</td>
<td>1.63 (1.09–2.43)</td>
<td>NA</td>
</tr>
<tr>
<td>Ronco et al. (2007) Italy</td>
<td>25–60</td>
<td>ASC-US</td>
<td>22,466</td>
<td>22,708</td>
<td>1.17 (0.87–1.56)</td>
<td>0.84 (0.56–1.25)</td>
</tr>
<tr>
<td>Ronco et al. (2007) Italy</td>
<td>25–60</td>
<td>LSIL</td>
<td>22,466</td>
<td>22,708</td>
<td>1.03 (0.74–1.43)</td>
<td>0.72 (0.46–1.13)</td>
</tr>
<tr>
<td>Strander et al. (2007) Sweden</td>
<td>23–60</td>
<td>ASC-US</td>
<td>8810</td>
<td>4674</td>
<td>1.40 (0.99–1.98)</td>
<td>NA</td>
</tr>
<tr>
<td>Maccallini et al. (2008) Italy</td>
<td>26–64</td>
<td>ASC-US</td>
<td>4299</td>
<td>4182</td>
<td>1.24 (0.72–2.15)</td>
<td>NA</td>
</tr>
<tr>
<td>Sykes et al. (2008) New Zealand</td>
<td>16–75</td>
<td>ASC-US</td>
<td>453</td>
<td>451</td>
<td>1.00 (0.83–1.21)</td>
<td>NA</td>
</tr>
<tr>
<td>Siebers et al. (2008) Netherlands</td>
<td>30–60</td>
<td>ASC-US</td>
<td>40,047</td>
<td>48,941</td>
<td>1.00 (0.84–1.20)</td>
<td>1.05 (0.86–1.29)</td>
</tr>
<tr>
<td>Siebers et al. (2009) Netherlands</td>
<td>30–60</td>
<td>ASC-US</td>
<td>40,047</td>
<td>48,941</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Klug et al. (2013) Germany</td>
<td>≥ 20</td>
<td>LSIL</td>
<td>9296</td>
<td>11,331</td>
<td>2.74 (1.66–4.53)</td>
<td>2.87 (1.55–5.32)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Reference Country</th>
<th>No. of women</th>
<th>Study design</th>
<th>Setting</th>
<th>Age at entry (years)</th>
<th>LBC procedure</th>
<th>Type of comparison</th>
<th>Reported end-points</th>
<th>Long-term outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blanks &amp; Kelly (2010)</strong> England</td>
<td>~2.5 million 102 laboratories, 13 643 abnormal tests</td>
<td>Before-and-after analysis of aggregated quality assurance data from screening laboratories</td>
<td>Organized screening</td>
<td>25–64</td>
<td>ThinPrep; SurePath</td>
<td>Before and after in laboratories that shifted from Conv. to LBC during 2005–2008</td>
<td>PPV</td>
<td>No</td>
</tr>
<tr>
<td><strong>Akamatsu et al. (2012)</strong> Japan</td>
<td>LBC: 29 119 Conv.: 49 108</td>
<td>Results for 2 consecutive rounds of screening during the shift from conventional Pap testing to LBC</td>
<td>Organized screening</td>
<td>NR</td>
<td>SurePath</td>
<td>Round 1: LBC vs Conv. Round 2: Conv. then Conv. Conv. then LBC LBC then LBC</td>
<td>Detection at round 1 and at round 2</td>
<td>Yes; CIN2+, CIN3+, and cervical cancer detection at next round</td>
</tr>
<tr>
<td><strong>Rebolj et al. (2015)</strong> Denmark</td>
<td>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</td>
<td>Before-and-after study with concomitant control</td>
<td>Organized screening</td>
<td>23–59</td>
<td>ThinPrep + ThinPrep Imaging System SurePath + FocalPoint + HPV triage for ASC-US</td>
<td>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</td>
<td>Relative referral Relative detection Relative PPV</td>
<td>No</td>
</tr>
<tr>
<td>Reference Country</td>
<td>No. of women</td>
<td>Study design</td>
<td>Setting</td>
<td>Age at entry (years)</td>
<td>LBC procedure</td>
<td>Type of comparison</td>
<td>Reported end-points</td>
<td>Long-term outcomes</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>---------</td>
<td>---------------------</td>
<td>---------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Rozemeijer et al. (2016, 2017) Netherlands</td>
<td>Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817</td>
<td>Concomitant comparison in cohort study; women may change the exposure over time</td>
<td>Organized screening</td>
<td>29–63</td>
<td>ThinPrep; SurePath</td>
<td>Conv. vs ThinPrep, Conv. vs SurePath, SurePath vs ThinPrep</td>
<td>Long-term outcome: incidence of cancers after negative screening test</td>
<td>Yes; cumulative incidence of cervical cancer</td>
</tr>
<tr>
<td>Ito et al. (2020) Japan</td>
<td>3 815 131 ThinPrep: 3 057 810 SurePath: 757 321</td>
<td>Concomitant comparison in cohort study; women may change the exposure over time</td>
<td>Spontaneous and organized screening</td>
<td>≥ 20</td>
<td>ThinPrep; SurePath</td>
<td>Conv. vs any LBC Poisson regression to compare adjusted detection of CIN2+ and CIN3+</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Reference, Country</th>
<th>No. of women</th>
<th>Referral for further assessment</th>
<th>Detection of CIN2+ and CIN3+</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blanks &amp; Kelly (2010)</strong> England</td>
<td>~2.5 million 102 laboratories, 13 643 abnormal tests SurePath and ThinPrep</td>
<td>NA</td>
<td>NA</td>
<td>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%</td>
</tr>
<tr>
<td>Akamatsu et al. (2012) Japan</td>
<td>Conv.: 49 108 LBC: 29 119 SurePath and ThinPrep</td>
<td>NA</td>
<td>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)</td>
<td>NA</td>
</tr>
<tr>
<td>Sigurdsson (2013) Iceland</td>
<td>103 909 Pap tests in 61 574 women 42 654 LBC tests in 20 439 women</td>
<td>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 &lt; 0.001) Women aged 40–69 yr: 0.88 (P that ratio is different from 1 = 0.026)</td>
<td>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)</td>
<td>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%</td>
</tr>
<tr>
<td>Reference Country</td>
<td>No. of women</td>
<td>Referral for further assessment</td>
<td>Detection of CIN2+ and CIN3+</td>
<td>PPV</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------</td>
<td>--------------------------------</td>
<td>----------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Rebolj et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>Conv. always: before, 47 300; after, 53 979</td>
<td>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)</td>
<td>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)</td>
<td>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)</td>
</tr>
<tr>
<td>Rozemeijer et al. (2016, 2017)</td>
<td>Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817</td>
<td>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)</td>
<td>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)</td>
<td>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)</td>
</tr>
<tr>
<td>Ito et al. (2020)</td>
<td>ThinPrep: 3 057 810 SurePath: 757 321</td>
<td>Conv.: 1.13% (34 435) LBC: 1.49% (11 443)</td>
<td>Adjusted RR, LBC vs Conv.: CIN2+: 1.16 (95% CI, 1.08–1.25) CIN3+: 1.00 (95% CI, 0.90–1.11)</td>
<td>Adjusted RR, LBC vs Conv.: CIN2+: 1.17 (95% CI, 1.09–1.26) CIN3+: 1.01 (95% CI, 0.91–1.12)</td>
</tr>
</tbody>
</table>

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; Bezrukow, 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

<table>
<thead>
<tr>
<th>Reference Country</th>
<th>Design</th>
<th>No. of women</th>
<th>Detection</th>
<th>IRR or RR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strander et al. (2007)</strong> Sweden</td>
<td>RCT with 3 yr and 7 mo follow-up</td>
<td>Conv.: 8810 LBC: 4674</td>
<td>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)</td>
<td>RR, 1.12 (95% CI, 0.68–1.83)</td>
</tr>
<tr>
<td><strong>Akamatsu et al. (2012)</strong> Japan</td>
<td>Results for 2 consecutive rounds of screening during the shift from conventional Pap testing to LBC</td>
<td>Conv. then Conv.: 73 253 Conv. then LBC: 33 318 LBC then LBC: 51 723</td>
<td>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</td>
<td>LBC then LBC vs Conv. then LBC: CIN2+: RR, 0.70 (95% CI, 0.44–1.11) CIN3+: RR, 0.64 (95% CI, 0.35–1.18) Cancer: RR, 0.32 (95% CI, 0.01–6.19)</td>
</tr>
<tr>
<td><strong>Rozemeijer et al. (2017)</strong> Netherlands</td>
<td>Concomitant comparison in cohort study; women may change the exposure over time</td>
<td>Conv.: 3 028 865 ThinPrep: 1 591 792 SurePath: 1 303 817</td>
<td>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</td>
<td>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)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Country</th>
<th>Position of official guidelines</th>
<th>Use of technology</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belarus</td>
<td>NA</td>
<td>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</td>
<td>IARC (2012)</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>NA</td>
<td>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</td>
<td>Aimagambetova &amp; Azizan (2018)</td>
</tr>
<tr>
<td>Republic of Moldova</td>
<td>Recommendation to progressively change from Romanowsky–Giemsa staining to Pap staining during the course of 2017</td>
<td>Opportunistic screening, with the majority using Romanowsky–Giemsa staining</td>
<td>Davies et al. (2016)</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>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</td>
<td>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.]</td>
<td>Olson et al. (2016); Ministry of Health of the Russian Federation (2019)</td>
</tr>
<tr>
<td>Soviet Union</td>
<td>Order No. 1253 (30 December 1976) introduced the use of Romanowsky–Giemsa staining across the whole country</td>
<td>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</td>
<td>Rogovskaya et al. (2013); Olson et al. (2016)</td>
</tr>
<tr>
<td>Turkmenistan</td>
<td>Order No. 144 (2014)</td>
<td>Pap testing followed by Romanowsky–Giemsa staining or HPV testing</td>
<td>WHO (2019)</td>
</tr>
</tbody>
</table>

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

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, Hoye 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 (£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

<table>
<thead>
<tr>
<th>Reference</th>
<th>Country</th>
<th>Date</th>
<th>Setting</th>
<th>Population/no. of tests (N)</th>
<th>Unsatisfactory samples</th>
<th>Detection rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iskhakova et al. (2012)</td>
<td>Russian Federation</td>
<td>2009–2011</td>
<td>Meleuz (Bashkortostan), centralized cytological laboratory</td>
<td>79,710 women, aged 20–60 yr</td>
<td>Unsatisfactory: 0.6%  Insufficiently satisfactory: 15.6% Satisfactory: 80.3%</td>
<td>CINI: 0.3 (n = 168) CINI: 0.2 (n = 86) CIN3: 0.05 (n = 43) CIS: 0.02 (n = 13) Cervical cancer: 0.02 (n = 13)</td>
</tr>
<tr>
<td>Kozyreva et al. (2012)</td>
<td>Russian Federation</td>
<td></td>
<td>Vladikavkaz (North Ossetia), oncological dispensary</td>
<td>9,525 nuclei of malignant and normal cervical cells</td>
<td>NR</td>
<td>Number detected (%): CINI: 530 (0.056%) CINI2: 960 (0.100%) CIN3: 890 (0.093%)</td>
</tr>
<tr>
<td>Chernyakova (2016)</td>
<td>Ukraine</td>
<td>2015</td>
<td>Kharkiv, university clinic</td>
<td>37 women aged 20–64 yr enrolled in “opportunistic screening”</td>
<td>19 women – inflammation; <em>Chlamydia</em> in 9.6%, HPV in 28.5% CINI in 2/37 Cytology–colposcopy discrepancy in 5/37</td>
<td>NR</td>
</tr>
<tr>
<td>Davies et al. (2016)</td>
<td>Republic of Moldova</td>
<td>2015</td>
<td>National audit Data from 7 of the largest laboratories</td>
<td>236,579 smears</td>
<td>Between laboratories, proportion of abnormal results varied from 0.32% to 6.06%, and unsatisfactory results varied from 0.0% to 5.7%</td>
<td>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</td>
</tr>
<tr>
<td>Aktanko et al. (2018)</td>
<td>Russian Federation</td>
<td></td>
<td>Vladivostok</td>
<td>4,032 women, aged &gt; 25 yr</td>
<td>NR</td>
<td>CINI: 21.9 (n = 20) CINI2: 12.1 (n = 11) CIN3: 19.7 (n = 18) CIS: 4.4 (n = 4) SCC: 30.7 (n = 28) Adenocarcinoma: 1.09 (n = 1)</td>
</tr>
<tr>
<td>Grebenkina et al. (2018)</td>
<td>Russian Federation</td>
<td>2018</td>
<td>Nizhny Novgorod, reference cytological centre; evaluated 10% of all cytological and 100% of all indeterminate samples</td>
<td>9,415 cytological smears 12% processed by Romanowsky–Giemsa</td>
<td>23% of all slides (not only Romanowsky–Giemsa stained smears)</td>
<td>21–36% did not match the final diagnosis (including 2 missed cervical cancers)</td>
</tr>
<tr>
<td>Reference</td>
<td>Country</td>
<td>Setting</td>
<td>Population/no. of tests (N)</td>
<td>Unsatisfactory samples</td>
<td>Detection rate (%)</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>--------------------------------</td>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>Kirillina et al. (2018)</td>
<td>Russian Federation</td>
<td>Yakutia, different women’s clinics</td>
<td>7600 women, aged 18–88 yr</td>
<td>Non-informative material: 1.9% Glandular epithelium not taken: 19.4%</td>
<td>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)</td>
<td></td>
</tr>
</tbody>
</table>

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.

(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


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.


Cervical cancer screening


Cervical cancer screening


